

```

CC or send an email to license@sb-sib.ch).
CC EMBL: D63346; BAA09670.1; -
DR EMBL: X73325; CAA51751.1; -
DR HSP; P01730; 1WBR.
DR GO; GO:0042101; C/T-cell receptor complex; ISS.
DR GO; GO:0015026; Fc receptor activity; ISS.
DR GO; GO:0042289; Fc receptor class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P-T-cell differentiation; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR000973; CD4 TCAG.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00447; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_Like; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 325
FT CARBOHYD 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 422
FT LIPID 422 458
FT CONFLICT 57 57
FT CONFLICT 91 91
FT CONFLICT 105 105
FT CONFLICT 113 113
FT CONFLICT 302 302
FT CONFLICT 349 349
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2C8B3E16 CRC64;

Query Match 33.4%; Score 903; DB 1; Length 458;
Best Local Similarity 87.0%; Pred. No. 1.3e-50;
Matches 174; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

```

```

DT 21-JUL-1986 (Rel. 01, Last sequence update)
DR 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig gamma chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RT F-I haplotype."
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RT immunoglobulin G of different allotype.",
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steimetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RT heavy chain and identification of two genomic C gamma genes.",
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RT immunoglobulin G."
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killester J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -1- MISCELLANEOUS: Ref.1 sequence has the D12 allotypic marker,
CC 104-Thr, and the E14 marker, 185-Thr. Ref.3 has the D11 and E15
CC markers and Ref.5 the E15 marker.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: M16426; AAA31289.1; -.
DR PIR: A91749; GHRA.
DR HSP; P01857; IFCL.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS50835; IG_Like; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 96
FT DOMAIN 114 213
FT DOMAIN 222 318
FT VARIANT 104 104
FT VARIANT 185 185
FT CONFLICT 48 48
FT CONFLICT 71 71

```

FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).  
 FT CONFLICT 173 173 N -> D (IN REF. 5).  
 FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).  
 FT CONFLICT 201 201 N -> D (IN REF. 5).  
 FT CONFLICT 218 218 Q -> E (IN REF. 5).  
 FT CONFLICT 233 233 E -> Q (IN REF. 5).  
 FT CONFLICT 246 246 N -> D (IN REF. 5).  
 FT CONFLICT 256 256 E -> G (IN REF. 5).  
 FT CONFLICT 260 260 N -> D (IN REF. 5).  
 FT CONFLICT 266 266 N -> D (IN REF. 5).  
 FT CONFLICT 280 280 Y -> W (IN REF. 5).  
 FT CONFLICT 284 284 N -> S (IN REF. 5).  
 SQ SEQUENCE 323 AA; 35404 MW; 6958AA118D579A8B CRC64;

Query Match 33.3%; Score 899; DB 1; Length 323;  
 Best Local Similarity 61.0%; Pred. No. 1.5e-50;  
 Matches 178; Conservative 34; Mismatches 64; Indels 16; Gaps 5;

QY 142 TLESPGSSPSVQCRPRGKNIQGGKTLVSQLELDGSGTWCTVYLQNKVFEKIDIVP 201  
 DB 45 TLTVNGVTFPSVQSS-----GLYSLSVSVSVTS SQPVTQNVVA--HPATNTKVD--- 92  
 QY 202 CPAPRPSKCDKTHTC--PELLGSPVFLPPPKKDTLMISTPEVTCVVNVVSHEDPEVK 259  
 DB 93 -KTVAPSTCSKP-TCPPPELLGSPSVFLPPPKKDTLMISTPEVTCVVNVVSHEDPEVK 150  
 QY 260 FNMVYDGVENNAKTPREEOYNSTRVSVYLVTHQDMLNGKEKCKVSNKALPAPIEK 319  
 DB 151 FTWYINNQVTRARPLAEQPNSTIRVSTLPIHQDMLNGKEKCKVSNKALPAPIEK 210  
 QY 320 TISKAKGPREPQVYTLPPSRDELTKNQVSLTCLVKGFPSPDIAVEMESNQEPENNYKT 379  
 DB 211 TISKARQPLEPKVYTMGPPEELSSRSVSLTCMNGFYPSDISVEMEKNGKAEDNYKT 270  
 QY 380 PVLVDSGSPFLYSKLTVDKSRMOQGNFSCSVHMEALHNHYTKSLSLSG 431  
 DB 271 PAVLDSDSYFLYNKLSVPTSEMQRGDVFTCSVMHEALHNHYTKSLSLSG 322

## RESULT 15

GC2\_CAVPO STANDARD; PRT; 329 AA.  
 AC P01862;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-2 chain C region.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.  
 OC NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE OF 1-3.  
 RA Trischmann T.M.;  
 RL Submitted (APR-1975) to the PIR data bank.  
 RN [2]  
 RP SEQUENCE OF 4-68.  
 RX MEDLINE=71058471; PubMed=5538606;  
 RA Birehstein B.K., Hussain Q.Z., Cebra J.J.;  
 RT "Structure of heavy chain from strain 13 guinea pig  
 RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the  
 RT half-cysteine joining heavy and light chains.";  
 RL Biochemistry 10:18-25(1971).  
 RN [3]  
 RP SEQUENCE OF 69-133 AND 312-329.  
 RX MEDLINE=71058486; PubMed=5538616;  
 RA Turner K.J., Cebra J.J.;  
 RT "Structure of heavy chain from strain 13 guinea pig  
 RT immunoglobulin-G(2). 11. Amino acid sequence of the carboxyl-terminal  
 RT and hinge region cyanoen bromide fragments.";  
 RL Biochemistry 10:9-17(1971).  
 RN [4]  
 RP SEQUENCE OF 134-226.

RX MEDLINE=75036072; PubMed=4429665;  
 RA Tracey D.E., Cebra J.J.;  
 RT "Primary structure of the CH2 homology region from guinea pig IgG2  
 RT antibodies.";  
 RL Biochemistry 13:4796-4803(1974).  
 RN [5]  
 RP SEQUENCE OF 227-311.  
 RX MEDLINE=75036073; PubMed=4609467;  
 RA Trischmann T.M., Cebra J.J.;  
 RT "Primary structure of the CH3 homology region from guinea pig IgG2  
 RT antibodies.";  
 RL Biochemistry 13:4804-4811(1974).  
 RN [6]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=71058474; PubMed=4922544;  
 RA Oliveira B., Lamm M.E.;  
 RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";  
 RL Biochemistry 10:26-31(1971).  
 CC -1- MISCELLANEOUS: This chain was isolated from pooled serum of strain  
 CC 13 inbred guinea pigs.  
 DR PIR: A94553; G2GP.  
 DR HSSP; P01842; 7RAB.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig\_c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SMO00407; IgC1; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; glycoprotein.  
 FT NOY TER 1  
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 28 79  
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 142 202  
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. .).  
 FT DISULFID 248 308  
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 33.1%; Score 894.5; DB 1; Length 329;  
 Best Local Similarity 62.5%; Pred. No. 2.9e-50;  
 Matches 172; Conservative 34; Mismatches 60; Indels 9; Gaps 4;

QY 163 IQGGKTLVSQLELDGSGTWCTVYLQ--NQKXVEFKIDIVCPAPRPSKCDKTHTC--PE 218  
 DB 58 LQSGLSYLTSMVTVVPSQKATCVNVAHPASSTKVDKVEPIRTPEBPCTCPK--CPPE 114  
 QY 219 LLAGPSVFLPPPKKDTLMISTPEVTCVVNVVSHEDPEVKFNMVYDGVENNAKTPRE 278  
 DB 115 NLGSPSVFIFPPPKKDTLMISTPEVTCVVNVVSHEDPEVQFTWVUNKVGNATPRV 174  
 QY 279 EQYNSTRVSVYLVTHQDMLNGKEKCKVSNKALPAPIEKTISKAKGPREPOVYTLPP 338  
 DB 175 EGYNTTFRVSIVLPIDHQDMLRGKEFKCKYNNKALPAPIEKTISKAGAPRMPVYTLPP 234  
 QY 339 SPBELTKNQVSLTCLVKGFPSPDIAVEMESNQGP--ENNYKTTPPVLDSDGSPFLYSKLT 396  
 DB 235 SRDLSSKSVSVTCLTINFPADIDHEWASNRPVSGKEKKNTPPIEDADSYFLYSKLT 294  
 QY 397 VDKSRMOQGNFSCSVHMEALHNHYTKSLSLSG 431  
 DB 295 VDKSAMQGTIVYTCVSHMEALHNHYTKSLSLSG 329

## RESULT 16

CD4\_CERAE STANDARD; PRT; 458 AA.  
 AC CD08338; O02805; O77593; Q28217;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen  
T4/Leu-3).  
GN CD4.  
OS Cercopithecus aethiops (Green monkey) (Grivet).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
RN NCBI\_TaxID=9534;  
RN [1]  
RA SEQUENCE FROM N.A.  
RA Hashimoto O., Tatemura M.;  
RT "Molecular cloning and expression of african green monkey CD4";  
RT Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 28-424 FROM N.A.  
RC TISSUE=Blood;  
RA MEDLINE=93049640; PubMed1425921;  
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
RT "Cloning and sequences of primate CD4 molecules; diversity of the  
RT cellular receptor for simian immunodeficiency virus/human  
RT immunodeficiency virus";  
RT Eur. J. Immunol. 22:2973-2981(1992).  
RN [3]  
RN SEQUENCE OF 28-424 FROM N.A.  
RP TISSUE=Peripheral blood;  
RX MEDLINE=98017879; PubMed=9379478;  
RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
RT "Relation between phylogeny of African green monkey CD4 genes and  
RT their respective simian immunodeficiency virus genes";  
RL J. Med. Primatol. 26:120-128(1997).  
RN [4]  
RP SEQUENCE OF 107-192 FROM N.A.  
RA MEDLINE=98320644; PubMed=9656488;  
RA Harris E.E., Disocell T.R.;  
RT "Nuclear gene trees and the phylogenetic relationships of the  
RT mangabeys (Primates: Papionini).";  
RL Mol. Biol. Evol. 15:892-900(1998).  
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
CC receptor interaction. May regulate T-cell activation.  
CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D86589; BAA13132.1; -;  
DR EMBL; X73332; CAA51748.1; -;  
DR EMBL; AF001226; AAB60873.1; -;  
DR EMBL; AF001228; AAB60873.1; -;  
DR EMBL; AF057380; AAC25124.1; -;  
DR HSSP; P01730; 1MIQ.  
DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
DR GO; GO:0015026; F:coreceptor activity; ISS.  
DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
DR GO; GO:0006955; P:immune response; ISS.  
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.  
DR GO; GO:0030217; P:T-cell differentiation; ISS.  
DR GO; GO:0045058; P:T-cell selection; ISS.  
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.  
DR InterPro; IPR000973; CD4\_TcRg.  
DR InterPro; IPR007110; IG-TcRg.  
DR InterPro; IPR003596; IG\_v.  
DR Pfam; PF00047; IG\_2.  
DR PRINTS; PR00692; CD4TCANTIGEN.  
DR SMART; SM00406; IGV; 1.

DR PROSITE; PS50835; IG LIKE; 1.  
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;  
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.  
FT SIGNAL 1 25  
FT CHAIN 1 25  
FT DOMAIN 26 458  
FT TRANSMEM 26 396  
FT DOMAIN 397 418  
FT DOMAIN 419 458  
FT DOMAIN 26 125  
FT DOMAIN 126 203  
FT DOMAIN 204 317  
FT DOMAIN 318 374  
FT CARBOHYD 42 42  
FT CARBOHYD 281 281  
FT CARBOHYD 296 296  
FT CARBOHYD 325 325  
FT DISULFID 41 109  
FT DISULFID 155 184  
FT DISULFID 328 370  
FT LIPID 419 419  
FT LIPID 422 422  
FT CONFLICT 46 46  
FT CONFLICT 59 59  
FT CONFLICT 115 115  
FT CONFLICT 165 165  
FT CONFLICT 200 200  
FT CONFLICT 227 227  
FT CONFLICT 271 271  
FT CONFLICT 281 281  
SQ SEQUENCE 458 AA; 51158 MW; FCS23D2EDD1F72E7 CRC64;  
Query Match 32.8%; Score 885; DB 1; Length 458;  
Beef Local Similarity 85.5%; Pred. No. 1.8e-49;  
Matches 171; Conservative 12; Mismatches 17; Indels 0; Gaps 0;  
QY 1 MNRGVPRHLLLVLTQALLPATQGNKVLLKGGDTVELTCTASQKSIQPFHKNNSNOIK 60  
DB 1 MNMGIPRHLVLVLTQALLPATVQKGVLLKGGDTVELTCTASQKTTQFMKNSNQIK 60  
QY 61 ILGNQSFVLTGPKSLNDRADSRSLMDQGNPPLIKLKIKEDSTTYICEVEDQKEVQL 120  
DB 61 ILGKQSFVLTGKSSSLRDRIDSRKSLMDQGFSMILKLIKEDSTTYICEVENKKEVEL 120  
QY 121 LVFGLTANSDPHLQGGSLTTLTLESPGSSPVQCRSRGNKIQGGKTLVSQLELOSG 180  
DB 121 LVFGLTANSDPHLQGGSLTTLTLESPGSSPVQCRSRGNKIQGGKTLVSQLELOSG 180  
QY 181 TWTCTVLQNKQKVERKIDIV 200  
DB 181 TWTCTVSDQDQNTVERKIDIM 200  
RESULT 17  
GCM MOUSE STANDARD; PRT; 405 AA.  
AC P01867;  
DT 21-VOL-1986 (Rel. 01, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE IG gamma-2b chain C region, membrane-bound form.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE OF 335-405 FROM N.A.  
RX MEDLINE=82222190; PubMed=6283537;  
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;  
RT "Nucleotide sequence of gamma chains.";  
RT Immunoglobulin gamma chains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
RN [2]  
RP SEQUENCE OF 335-378 FROM N.A.

```

RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
CC
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Membrane-bound;
CC IsoId=PI01867-1; Sequence=Displayed;
CC Name=Secreted;
CC IsoId=PI01866-1; Sequence=External;
CC Note=May be the major isoform;
CC -1- PTM: DISULFIDE BONDS BY HOMOLOG WITH OTHER IG GAMMA CHAINS.
CC -1- MISCELLANEOUS: The sequence of residues 1-335 is assumed to be
CC identical with the corresponding region of the secreted form.
CC -1- MISCELLANEOUS: The a allele sequence is shown.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, J00462; AAB59659.1; ALT_INIT.
CC PIR, C02154; G2MSBM.
CC PDB, 1C1C; 1I-MAR-03.
CC MGD; MGI:96445; Igh-3.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003597; IG_c1.
CC InterPro: IPR003006; IG_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IgC1; 2.
CC PROSITE; PS50835; IG_LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Transmembrane;
CC Alternative splicing; 3D-structure; Repeat.
CC
CC FT NON_TER 1
CC FT DOMAIN 6
CC FT DOMAIN 127
CC FT DOMAIN 225
CC FT DISULFID 15
CC FT DISULFID 27
CC FT DISULFID 109
CC FT DISULFID 112
CC FT DISULFID 115
CC FT DISULFID 118
CC FT DISULFID 150
CC FT DISULFID 210
CC FT DISULFID 256
CC FT TRANSMEM 352
CC FT DOMAIN 370
CC SQ SEQUENCE 405 AA; 44330 MW; 89B3CF0A9B6D49FA CRC64;

Query Match 32.5%; Score 877.5; DB 1; Length 405;
Best Local Similarity 51.7%; Pred. No. 4.7e-49;
Matches 178; Conservative 52; Mismatches 75; Indels 39; Gaps 9;

OY 133 LIQG--OSLTLTLESPPGSS-----PSVQCRSPRGKNIQGGKTLVSQLELDQSGTW-- 182
DB LVKGYFPESVTVTNGSLSSSVHTFPAL-----LQSG-LVTMSSSVTVPSSTWPS 77
OY 133 ---TCTVQ--NQKVEFKID-----IVCPAPPEPSCKDTHTC--PELLGSPSVLF 228
DB QTVTSVAHPASSTVDDKLEPSGPISTINPCP-----PCKECHCPAPNLEGGSVFIF 132
OY 229 PPKKDTLMISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVNAKTKRREQVSTYVYV 288
DB 133 PPNIDVLMISTLPKVTCCVVVDVSDDDPVQVLSMFWNVNVEVHTATQTHREDVNSTIRVV 192
OY 289 SVLTVLHODWINGEKYKCKVSNKALPAPIEKTIISKAKQPREPQVYTLPPSRDELTKNOV 348

```

```

DB 193 STLFIQHODWMSGKFEKCKVNNKXLPSPIBRTISKIGLVAPQVILLPPAEQLSRKDV 252
OY 349 SLTCLVGFYPSDAVAMENSGOPENNYKTPPLVDSGSEFLYSKTLVDSRQGVNF 408
DB 253 SLTCLVGFNPGDLSVMTSGNTEENYKOTAPVLDSDGSFYISKNNMKTISKWEKTDSE 312
OY 409 SCVYHEBALNHYTKSLSLSPGLDPECAEADGSLDGLMTT 452
DB 313 SCVYHGLKKNYIKTKTISRSPLGLDDICAEADGSLDGLMTT 356

RESULT 18
ID3_MOUSE STANDARD; PRT; 329 AA.
AC P22436.
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=65027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blatter F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBL J. 3:2041-2046(1984).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, J00451; -; NOT_ANNOTATED_CDS.
CC PIR, B02156; G3MSC.
CC HSSP; P01857; 1PCL1.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003597; IG_c1.
CC InterPro: IPR003006; IG_MHC.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; IgC1; 2.
CC PROSITE; PS50835; IG_LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
CC Transmembrane; Alternative splicing.
CC
CC FT NON_TER 1
CC FT DOMAIN 1
CC FT DOMAIN 98
CC FT DOMAIN 114
CC FT DOMAIN 224
CC SQ SEQUENCE 329 AA; 36628 MW; F45827174182BAD6 CRC64;

Query Match 31.7%; Score 857.5; DB 1; Length 329;
Best Local Similarity 59.6%; Pred. No. 6.8e-48;
Matches 162; Conservative 38; Mismatches 57; Indels 15; Gaps 4;

OY 170 SVSGLLELDQSGTW-----TCTVLNQKVEFKID--VCPAPPEPSCKDTHTC--EL 219
DB 62 SLSSLVTVPSSTWPSQTVICNVAPASKTELIRKIEIRIKPSPSPS-----SCPGNI 116
OY 220 LGGSPVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVNAKTKRREE 279
DB 117 LGGSPVFIFFPKKDTLMISTLPKVTCCVVVDVSDDDPVQVLSMFWNVNVEVHTATQTHREA 176
OY 280 QNSTYRVSVLTVLHODWINGEKYKCKVSNKALPAPIEKTIISKAKQPREPQVYTLPPS 339

```



```
Db 177 QVNSTFRVVSALPIQHODMMRCKEFCVKVNNKALPAPIERTISKPKRAQTPQVYTIPTPP 236
Qy 340 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK 399
Db 237 RQMSGKKRSLTCLVNFSEALSVEMERNGELQDYKNTTPPLDSGTYFLYSKLTVDY 296
Qy 400 SRMQGQVSCSVMEALHNHYTQKSLSLSPG 431
Db 297 DSWLQGEIFTCVSVEALHNHHTQKLSRSPG 328

RESULT 19
GCB_RAT ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Bruggemann M.;
RL "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig1; 3.
DR SMART; SM00407; IG1; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 96 IG-LIKE 1.
FT DOMAIN 124 223 IG-LIKE 2.
FT DOMAIN 232 328 IG-LIKE 3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 30.8%; Score 833; DB 1; Length 333;
Best Local Similarity 58.6%; Pred. No. 2, 5e-46;
Matches 157; Conservative 38; Mismatches 53; Indels 20; Gaps 5;

179 SGTW-----TCTVLAQ--NOKVVEFKIDI-----VCPAPEPKSCDKHTTC--PELLGSP 223
Db 70 SSTWPSQATYTCNVAVHAPASSTKVDKVERNRNGIGHKCP-----TCTGCHKCPVELLGGP 124
Qy 224 SYFLPPPKKDTLMISTREYTCVVVDVSHEDPEVKFNWYVDGVEYNNAKTKPREEQYNYS 283
Db 125 SVFIPEPKPKDILLISQNAKVTCTVVDVSEEDVQSFVFNWVEVHTAQOTPREEQYNYS 184
Qy 284 TRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTK 343
Db 185 TRRVVSALPIQHODMMRCKEFCVKVNNKALPAPIEKTISKPKRAQTPQVYTIPTPP 244
Qy 344 TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRMQ 403
Db 245 TQGVSLTCLTSGFLNDIGEVWTSNGHTEKVKNTPEVMDSDGSFFMYSKLNVERSRD 304
```

```
Qy 404 QGNVFCSVMEALHNHYTQKSLSLSPG 431
Db 305 SRAPVCSVMEGLHNHVKISIRPPG 332

RESULT 20
GCL_RAT ID GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-1 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Bruggemann M.;
RL "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC PIR; PS0017; PS0017.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig1; 3.
DR SMART; SM00407; IG1; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 30.5%; Score 823.5; DB 1; Length 326;
Best Local Similarity 50.0%; Pred. No. 1e-45;
Matches 163; Conservative 53; Mismatches 55; Indels 55; Gaps 9;

133 ILQG---QSLTTLTSPGSS-----PSVQCRSPGKNIQGGKTLVSQLELDQSGTW-- 182
Db 28 LVKGFPEPVYTNMGSALSSGVTFRPVA-----LQSLTYLTSSVTV--PSTWMS 77
Qy 183 ---TCTVLAQ--NOKVVEFKIDI--VCPAPEPKSCDKHTTCPELLG-----PSV 225
Db 78 QTVTCNVAVHAPASSTKVDKDI-----VPRNC-----GDDCKPCICTGSEVSVS 119
Qy 226 FLFPPPKKDTLMISTREYTCVVVDVSHEDPEVKFNWYVDGVEYNNAKTKPREEQYNYS 285
Db 120 FLFPPPKKDVLTTLTPKTCVVVDISQDDPVHSEWFDVDEVHTAQOTPREEQYFNSTP 179
Qy 286 RVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTK 345
Db 180 RVSSELPIHQMDMLNGKTRFCVTSAAFPSPLEKTIISKPEGTQVPHYTNMSPTEBMTQ 239
Qy 346 NOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRMQ 405
Db 240 NEVSITCWKGFYPSDIAVEWESNGQPOBNYKNTPEVMDSDGSFFLYSKLTVDKSRMQ 299
```

Oy 406 NVFSCSVNHEALHNHYTKSLSPG 431  
 Db 300 NTFTCSVLHEGLHNHTEKSLSHSPG 325

## RESULT 21

GCC\_RAT STANDARD; PRT; 329 AA.  
 ID\_GCC\_RAT  
 AC P20762;  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig gamma-2C chain C region.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NC NCB1\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88166903; PubMed=3127222;  
 RA Brueggemann M., Delmastro-Galife P., Waldmann H., Calabi F.;  
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant  
 RT region cDNA: extensive homology to mouse gamma 3.";  
 RT Eur. J. Immunol. 18:317-319(1988).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X07189; CAA30169.1; -  
 DR PIR: S00847; S00847.  
 DR HSR: F01842; 7PAB.  
 DR InterPro: IPR007110; IG-1like.  
 DR InterPro: IPR003597; IG\_C1.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR Pfam: PF00047; Ig; 2  
 DR SMART: SM00407; IGc1; 2  
 DR PROSITE: PS50835; IG\_LIKE; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KM Immunoglobulin domain; Immunoglobulin C region.  
 FT NON TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 222 CH2.  
 FT DOMAIN 223 329 CH3.  
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 82  
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 143 203  
 FT DISULFID 249 307  
 SQ SEQUENCE 329 AA; 36571 MW; 5PCD7B7933850773 CRC64;

Query Match 30.5%; Score 823; DB 1; Length 329;  
 Best Local Similarity 58.5%; Pred. No. 1.1e-45;  
 Matches 151; Conservative 45; Mismatches 56; Indels 6; Gaps 2;

Oy 179 SGTW-----TCTVLNOKKVEFKIDIVCPAPEPKSCDKHTTCEPLLGSPVFLPFPKPK 233  
 Db 72 SSTWSQVTCGVAPATKSNLIKRIE-RRKKRPPTDICSQDNLRSPVFTFPKPK 130  
 Oy 234 DTLMSIRPEVTCVVVDVSHEDPEVKFMVYDGVENNAKTKPREQYNSTYRVVSVLTV 293  
 Db 131 DILMITLTPKVTQVVVDVSEEDPDVQFSPFDNVAVFTAQNPHEQNGFRVYSTLHI 190  
 Oy 294 LHDQNLNKEYKCKSNKALPAPIKETISKAKGQREGEVYTLTPSRBELTKNQVSLTCL 353  
 Db 191 QHDDMSGKEPKCKVNNKDLPSPIKTIKSKPKARTQVYITIPPREQMSKNKSLTCLM 250

Oy 354 VKGFYPDIADVEMESNCPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOGNVSCSV 413  
 Db 251 VTSFYPASISVEMERNELQDYKNTLPVLDSDGSPFLYSKLTVDTSNRKGRDITCSV 310  
 Oy 414 HEALHNHYTKSLSPG 431  
 Db 311 HEALHNHYTKSLSPG 328

## RESULT 22

GCL\_MOUSE STANDARD; PRT; 324 AA.  
 ID\_GCL\_MOUSE  
 AC P01868;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ig gamma-1 chain C region secreted form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;  
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 RT gamma 1 chain gene.";  
 RL Cell 18:559-568(1979).  
 RN [2]  
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RX MEDLINE=8020559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 RT cloned in a bacterial plasmid.";  
 RL Gene 9:87-97(1980).  
 RN [3]  
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=80012837; PubMed=113776;  
 RA Rogers J., Clarke P., Salsber W.;  
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
 RT heavy chain.";  
 RL Nucleic Acids Res. 6:3305-3321(1979).  
 RN [4]  
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=78242288; PubMed=98524;  
 RA Adegbo K.;  
 RT "Evolution of immunoglobulin subclasses. Primary structure of a  
 RT murine myeloma gamma1 chain.";  
 RL J. Biol. Chem. 253:6068-6075(1978).  
 RN [5]  
 RP DISULFIDE BONDS (MOPC 21).  
 RX MEDLINE=73008889; PubMed=5073237;  
 RA Svasil J., Malsstein C.;  
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";  
 RL Biochem. J. 126:837-850(1972).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Secreted;  
 CC IsoId=P01868-1; Sequence=Displayed;  
 CC Note=May be the major isoform;  
 CC Name=Membrane-bound;  
 CC IsoId=P01869-1; Sequence=External;

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

CC -----
DR EMBL, V00793; CAA24172.1; -
DR EMBL, V00793; CAA24173.1; -
DR EMBL, V00793; CAA24174.1; -
DR EMBL, V00793; CAA24175.1; -
DR EMBL, V00795; CAA24176.1; -
DR PIR, A02159; GIMS.
DR GlycosylatedB; P01868; -.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR007110; Ig-1Ike.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 118 174
FT CARBOHYD 174 174
FT FTID=CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 30.3%; Score 818.5; DB 1; Length 324;
Best Local Similarity 51.8%; Pred. No. 2.1e-45;
Matches 156; Conservative 50; Mismatches 60; Indels 35; Gaps 5;

OY 133 LLAGGSLTLTLSPSGSSPSVOCSPRGNIOGKTLVSQLELDPSGTWCTVLQ--NQ 190
Db 56 VLQSLTLTLSSSVTVPSSP-----KPSVITCNVAHPSS 90
OY 191 KAVRKIDIVPCPAPEPKSCDKTHCPCLGSPVFLFPKPKDTLMSRPEVTCVVD 250
Db 91 TKVDKKIVPRDCGC-KPCIC---TVPEV---SSVFIFPKPKDVLITLTPKVCVVD 142
OY 251 VSHDEPEVKFNMYVGVGVHNAKTKPREQVNSTRVVSVLTVLHODMNGEKVKCYSN 310
Db 143 ISKDDPEVQFSMFVDVDEVHTAQTQREQFNSTRSVSELPIMHODMNGEKPEKRVNS 202
OY 311 KALPAPIEKTSKAGQPREPOVYTLPPSRDELTKQVSLTCLVAGFYPSDIIVEMESNG 370
Db 203 AAFPAPIEKTIKTKGRPAKAPQVYTLPPKKEGMADKDKSLTCLMIDFFPEDITVEMQNG 262
OY 371 QPENNYKTPPYLVDSDGSEFLYSKLTVDKSRNQGVNCSVMHEALNNHYTQKSLSLSP 430
Db 263 QPAENYKTPQPMNTNGSYFYVSKLNVQKSNWEAGNTFTCSYLAHGLNHNHTEKSLSHSP 322
OY 431 G 431
Db 323 G 323

```

```

DE Ig gamma-2A chain C region, A allele.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1)
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RN Nucleic Acids Res. 8:3143-3155(1980).
(2)
RP SEQUENCE FROM N.A.
RX MEDLINE=81189976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
RN Nucleic Acids Res. 9:1365-1381(1981).
(3)
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Olio R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
suggests that exons can be exchanged between genes in a multigenic
family.";
RN Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
(4)
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
immunoglobulin: amino-acid sequence of the Fc fragment. Implications
for the evolution of immunoglobulin structure and function.";
RN Eur. J. Biochem. 43:423-435(1974).
(5)
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC - SIMILARITY: Contains 3 immunoglobulin-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, V00798; CAA24178.1; -.
DR PIR, A02152; G2MSA.
DR PDB; 1B4W; 12-JUL-01.
DR PDB; 1B4X; 12-JUL-01.
DR PDB; 1MN0; 06-MAY-99.
DR InterPro; IPR007110; Ig-1Ike.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KM Immunoglobulin domain; Immunoglobulin C region; 3D-structure; Repeat.
FT NON_TER 1
FT DOMAIN 1 98 IG-LIKE 1.
FT DOMAIN 121 220 IG-LIKE 2.
FT DOMAIN 229 325 IG-LIKE 3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).

```

FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT MOD RES 330 330  
 SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 29.8%; Score 805.5; DB 1; Length 330;  
 Best Local Similarity 58.5%; Pred. No. 1.4e-44;  
 Matches 155; Conservative 33; Mismatches 58; Indels 19; Gaps 5;

QY 179 SGTW-----TCTVLO--NOKKVEKID-----IVCPAPPEKSCDKTTCBELLGGEPSVF 226  
 DB 72 SSTWPSOSITCVVAHPASTKVKDKIEPRGPITKPCP---PCKC-----PANNLLGGEPSVF 124  
 QY 227 LPPPKKDTLMSRPREYTCVAVDVSHDEPKFKNWYDVGVYVNAHAKTKPREEQNSYR 286  
 DB 125 IFPPKIKVLMISLSPVITCVAVDVSEDDPVQISFVNNVVEVHTAQOTRHEDNSTLR 184  
 QY 287 VSVVLTVLHODMLNGKEYCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSRDELTKN 346  
 DB 185 VVSALPIQHODMWSKEFKCKVNNKDLPAPIERTISKSGSVRAQVYVLEPPEEEMTKK 244  
 QY 347 QVSLTCLVKGFPSPDIAMVESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRWQGN 406  
 DB 245 QVTLTCMTDFWPEPDIYVEMTNGKTELINXKTEFVLDSGSYFWYSKLRVKKWVERN 304  
 QY 407 VFSGSVMEALHNHYTOKSLSPG 431  
 DB 305 SYSCSVMEGLHNHTTYSFRTPG 329

## RESULT 24

GCAB\_MOUSE STANDARD; PRT; 335 AA.

AC P01864;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ig gamma-2A chain C region secreted form (B allele).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=82037861, PubMed=6170065;  
 RA Schreier P.H., Bothwell A.L.M., Mueller-Hill B., Baltimore D.;  
 RT "Multiple differences between the nucleic acid sequences of the  
 RT IgG2a and IgG2b alleles of the mouse.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).  
 RN [2]  
 RP SEQUENCE.  
 RX MEDLINE=82037777, PubMed=6794027;  
 RA Dognin M.J., Lauwereys M., Strosberg A.D.;  
 RT "Multiple amino acid substitutions between murine gamma 2a heavy  
 RT chain Cc regions of Ig1a and Ig1b allotypic forms.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Secreted;  
 CC IsoId=P01864-1; Sequence=Displayed;  
 CC Note=Probably the major isoform;  
 CC Name=Membrane-bound;  
 CC IsoId=P01865-1; Sequence=External;  
 CC -1- MISCELLANEOUS: The sequence differs from that of the a allele,  
 CC from BALB/c mice, at 15% of the positions.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----  
 CC EMBL; J00479; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A02153; G2MSAB.  
 DR PDB; 1BOG; 23-JAN-99.  
 DR PDB; 1HH6; 26-JAN-01.  
 DR PDB; 1HH9; 24-JUL-03.  
 DR PDB; 1HH6; 08-FEB-01.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig c1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig\_2.  
 DR SMART; SM00407; IgC1; 2.  
 DR PROSITE; PS50835; Ig\_LIKE; 3.  
 DR PROSITE; PS00290; Ig\_MHC; 1.  
 DR Immunoglobulin domain; Immunoglobulin C region; Alternative splicing;  
 DR 3D-structure; Repeat.  
 KW NON TER 1  
 FT DOMAIN 6 98 IG-LIKE 1.  
 FT DOMAIN 126 225 IG-LIKE 2.  
 FT DOMAIN 234 330 IG-LIKE 3.  
 SQ SEQUENCE 335 AA; 36596 MW; FA3382792CB813C6 CRC64;

## Query Match

29.6%; Score 801; DB 1; Length 335;  
 Best Local Similarity 50.2%; Pred. No. 2.8e-44;  
 Matches 166; Conservative 52; Mismatches 77; Indels 36; Gaps 9;

QY 124 GLTANSDT--HLQG---QSLTTLLESPGSS-----PSVQCRSPRGKNIQSGKTLVSQ 173  
 DB 17 GTTGSSVTLGCLVGYPEPVTLTWNSGSLSGVHTPAL-----LQSG-LYTLSS 66  
 QY 174 LELDQSGTW-----TCTV-----LQNKVVEKIDIV--PCAPPEKSCDKTTCBELL 220  
 DB 67 SVTLSTWTSQITTCVVAHPASTKVKDKIEPRVPTQWPCP---PQQRVPCAADLL 123  
 QY 221 GGPSEVFPFPKPKDTLMSRPREYTCVAVDVSHDEPKFKNWYDVGVYVNAHAKTKPREEQ 280  
 DB 124 GGPSEVFPFPKPKDTLMSRPREYTCVAVDVSEDDPVQISFVNNVVEVHTAQOTRHED 183  
 QY 281 YNSTYRVSVLTVLHODMLNGKEYCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSR 340  
 DB 184 YNSTLRVVSALPIQHODMWSKEFKCKVNNKDLPAPIERTISKSGSVRAQVYVLEPPE 243  
 QY 341 DELTKNOVSLTCLVKGFPSPDIAMVESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKS 400  
 DB 244 EEMTKKEFSLTCMITGFLPAEIAVDWTSNGRTEQNYKNTATVLDSDGSYFWYSKLRVQKS 303  
 QY 401 RMQGSNVFSGSVMEALHNHYTOKSLSPG 431  
 DB 304 TWERKSLFACSVMEGLHNHTTYSFRTSG 334

## RESULT 25

GCA\_RAT STANDARD; PRT; 322 AA.

AC P20760;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-2A chain C region.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89232738; PubMed=3149946;  
 RA Brueggemann M.;  
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";

```

RL Gene 74:473-482(1988).
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M13804; AAA41376.1; ALU_INIT.
CC PIR: PS0019; PS0019.
CC HSSP: P01842; 7EAB.
CC InterPro: IPR007110; Ig_Like.
CC InterPro: IPR003587; Ig_C1.
CC InterPro: IPR003006; Ig_MHC.
CC Pfam: PF00047; Ig_2.
CC SMART: SM00407; IgC1; 2.
CC PROSITE: PS00835; IG_LIKE; 3.
CC PROSITE: PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Repeat.
CC NON_TER 1 1
CC DOMAIN 6 98 IG-LIKE 1.
CC DOMAIN 115 212 IG-LIKE 2.
CC DOMAIN 221 317 IG-LIKE 3.
CC DISULFID 27 82
CC DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 136 196
CC DISULFID 242 300
CC CARBOHYD 172 172
CC SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;
SQ
Query Match 29.3%; Score 792.5; DB 1; Length 322;
Best Local Similarity 51.0%; Pred. No. 9,4e-44;
Matches 160; Conservative 47; Mismatches 72; Indels 35; Gaps 8;
OY 133 LIQG--QSLTLTLESPSS-----PSVQCSPPKNGIKQGGKTLSSVQLQDSGT-- 182
DB 28 LVYGPPEPVTVYMGALSSGVHTPPAV-----LQSGLYTLTSSVTV-PSSTWSS 77
OY 183 ----TCTVLQ--NOKKVEFKIDIVPCPAPRPSKCDKHTHCBELLGGSVSLFPKPKDTLM 237
DB 78 QAVTCAVAHPASTKVDKKIVPREC---NPCGTGSEV-----SSVPIFPKPTDVLTL 127
OY 238 IARTPEVTGVVDSHEDPEVKNVYVDGVEVNAATKPREQYNSTYRVVSVLTVLHOD 297
DB 128 ITLTPKVTGVVVDISQNDPEVRSMVIDVEVHTAQTTHAPRKQSNSTLRSVSELPVHRD 187
OY 298 WLNKGEKCKVSNKALPARIKTIKSAKQPREPOVYTLPPSHDELTKNQVSLTCLVKGF 357
DB 188 WLNKGTFFKCKVNSGAPFAPIKSIKRPETGPRGQYVYTAAPREEMTQGSVITCWVGF 247
OY 358 YPSDIAVESNGQPENNYKTPPVLDSDGSFYLKLTVDKSRKQGVNFGSVWHEAL 417
DB 248 YEPDITWKWNGQPENNYKTPPTDSDGSFYLKLVKWKETWQGNFTCSVLAEGL 307
OY 418 HNHVYTKSLSPG 431
DB 308 HNHHTKSLSHSPG 321

```

```

GN CD4.
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=95311;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus";
RL Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X73328; CAAS1754.1; -.
CC EMBL: X73327; CAAS1753.1; -.
CC HSSP: P01730; 1WTO.
CC DR GO: GO:0042101; C: T-cell receptor complex; ISS.
CC DR GO: GO:0015026; F: coreceptor activity; ISS.
CC DR GO: GO:0042289; F: MHC class II protein binding; ISS.
CC DR GO: GO:0006955; P: immune response; ISS.
CC DR GO: GO:0045086; P: positive regulation of interleukin-2 biosyn. .; ISS.
CC DR GO: GO:0030217; P: T-cell differentiation; ISS.
CC DR GO: GO:0045058; P: T-cell selection; ISS.
CC DR GO: GO:0007169; P: transmembrane receptor protein tyrosine kin. .; ISS.
CC DR InterPro: IPR000973; CD4 TCRG.
CC DR InterPro: IPR007110; Ig_Like.
CC DR InterPro: IPR003596; Ig_V.
CC DR Pfam: PF00047; Ig_2.
CC DR PRINTS: PR00692; CD4TCANTIGEN.
CC DR SMART: SM00406; IG_V; 1.
CC DR PROSITE: PS00835; IG_LIKE; 1.
CC DR Immunoglobulin domain; Glycoprotein; T-cell;
CC Immune response; Repeat; Lipoprotein; Palmitate.
CC NON_TER 1 1
CC DOMAIN <1 369 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 370 391 POTENTIAL.
CC DOMAIN 392 >397 CYTOPLASMIC (POTENTIAL).
CC DOMAIN <1 98 IG-LIKE V-TYPE.
CC DOMAIN 99 176 IG-LIKE C2-TYPE 1.
CC DOMAIN 177 290 IG-LIKE C2-TYPE 2.
CC DOMAIN 291 347 IG-LIKE C2-TYPE 3.
CC CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC DISULFID 14 82 BY SIMILARITY.
CC DISULFID 128 157 BY SIMILARITY.
CC DISULFID 301 343 BY SIMILARITY.
CC LIPID 392 392 S-palmitoyl cysteine (By similarity).
CC LIPID 395 395 S-palmitoyl cysteine (By similarity).
CC VARIANT 20 20 MISSING.
CC VARIANT 43 43 T -> I.
CC VARIANT 86 86 N -> D.
CC VARIANT 96 96 F -> L.
CC VARIANT 173 173 V -> M.
CC VARIANT 316 316 R -> K.
CC NON_TER 397 397

```

```

SQ SEQUENCE 397 AA; 43926 MW; 8660B636D2DB38A7 CRC64;
Query Match 29.0%; Score 784; DB 1; Length 397;
Best Local Similarity 87.3%; Pred. No. 4.3e-43;
Matches 151; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 28 VLGGKGDVTELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRSLW 87
DB 1 VLGGKGDVTELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRSLW 60
QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEEVOVLVFGLTANSPTHLLOGOSLTLTLESPP 147
DB 61 DQGCFSMIINKLKIEDSEITYICEVENKKEVEVLVFGLTANSPTHLLOGOSLTLTLESPP 120
QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTCTVTLONQKVEFKIDIV 200
DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTCTVTLONQKVEFKIDIV 173

RESULT 27
CD4_ERYPA STANDARD; PRT; 397 AA.
AC 008339; 1995 (Rel. 31, Created)
DT 01-EB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)
DE (Fragment).
OS Erythrocybus pates (Red guenon) (Husaei).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopitheciinae; Erythrocybus.
OX NCBI_TaxID=9538;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus."
RL Eur. J. Immunol. 22:2973-2981(1992).
CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC - SUBUNIT: Associates with p56-lck (By similarity).
CC - SUCCESSESSOR LOCATION: Type I membrane protein.
CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X73324; CAAS1750.1; -.
DR HSSP; P01730; 1M10.
DR GO; GO:0042101; C-T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR000973; CD4 TAG.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TANTIGEN.

```

---

```

DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW immune response; Repeat; Lipoprotein; Palmitate.
FT NON TER 1
FT DOMAIN 1
FT TRANSMEM 370
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 392
FT DOMAIN >397
FT DOMAIN <1
FT DOMAIN 98
FT DOMAIN 99
FT DOMAIN 176
FT DOMAIN 177
FT DOMAIN 290
FT CARBOHYD 291
FT CARBOHYD 269
FT CARBOHYD 298
FT DISULFID 14
FT DISULFID 82
FT DISULFID 128
FT DISULFID 157
FT DISULFID 301
FT LIPID 343
FT LIPID 392
FT LIPID 392
FT LIPID 395
FT LIPID 395
FT NON TER 397
SQ SEQUENCE 397 AA; 44081 MW; 67887397A6B7EA4F CRC64;
Query Match 29.0%; Score 783; DB 1; Length 397;
Best Local Similarity 86.7%; Pred. No. 4.9e-43;
Matches 150; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 28 VLGGKGDVTELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRSLW 87
DB 1 VLGGKGDVTELTCTASQKKSIOFHWKNSNQIKILGNQGSFLTQKPSKLNDRADSRSLW 60
QY 88 DQGNFPLIINKLKIEDSDTYICEVEDQKEEVOVLVFGLTANSPTHLLOGOSLTLTLESPP 147
DB 61 DQGCFSMIINKLKIEDSEITYICEVEDKKEVEVLVFGLTANSPTHLLOGOSLTLTLESPP 120
QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTCTVTLONQKVEFKIDIV 200
DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTCTVTLONQKVEFKIDIV 173

RESULT 28
GCB_MOUSE STANDARD; PRT; 336 AA.
ID GCB_MOUSE
AC P01866;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ig gamma-2B chain C region secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ALIELE A).
RC MEDLINE=80120716; PubMed=676534;
RX Yamawaki-Kataoka Y., Kataoka T., Takahashi N., Oba M., Honjo T.;
RA "Structure of the constant and 3' untranslated regions of the murine
RT gamma 2b heavy chain messenger RNA."
RL Science 206:1299-1303(1979).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=80081502; PubMed=117549;
RX Tucker P.W., Marcu K.B., Slichtom J.L., Blattner F.R.;
RA "Structure of the cloned gene for the constant region of murine gamma
RT 2b immunoglobulin heavy chain."
RL Science 206:1303-1306(1979).
RN [4]

```

```

RP SEQUENCE FROM N.A. (ALLELE B).
RX MEDLINE=82173203; PubMed=6803173;
RA Ollé R., Rougeon F.;
RT "Mouse immunoglobulin allotypes: post-duplication divergence of gamma
RL 2a and gamma 2b chain genes."
RL Nature 296:761-763(1982).
[5]
RN CARBOHYDRATE-LINKAGE SITE THR-105.
RX MEDLINE=94216359; PubMed=7512967;
RA Kim H., Yamaguchi Y., Masuda K., Matsunaga C., Yamamoto K.,
RA Irimura T., Ikehashi N., Kato K., Arata Y.;
RT "O-glycosylation in hinge region of mouse immunoglobulin G2b."
RL J. Biol. Chem. 269:12345-12350(1994).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential1).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoform=2;
CC Name=Secreted;
CC IsoId=P01866-1; Sequence=Displayed;
CC Note=May be the major isoform;
CC Name=Membrane-bound;
CC IsoId=P01867-1; Sequence=External;
CC MODIFIED WITH 2 SIALIC ACID RESIDUES.
CC -1- PTM: O-LINKED GLYCAN CONSISTS OF GAL-GALNAC DISACCHARIDE WHICH IS
CC -1- MISCELLANEOUS: BONDS BY HOMOLOGY WITH OTHER IG GAMMA CHAINS.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR PIR: S25057; G2MS11.
DR HSBP, P01842; 7FAB.
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00407; IGc1. 2.
DR PROSITE: PS50835; IG_LIKE; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KM Alternative splicing; Repeat.
FT NON TER 1 1
FT DOMAIN 1 98
FT 127 IG-LIKE 1.
FT 226 IG-LIKE 2.
FT 235 IG-LIKE 3.
FT 331 INTERCHAIN (WITH A LIGHT CHAIN).
FT 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT 118 118 INTERCHAIN (WITH A HEAVY CHAIN).
FT 210 210 INTERCHAIN (WITH A HEAVY CHAIN).
FT 256 314
FT 314
FT 105 105 O-LINKED (GALNAC...),
FT 336 REMOVED POST-TRANSLATIONALLY (PROBABLE).
FT 163 163 Q -> R (IN ALLELE B).
FT 194 194 T -> A (IN ALLELE B).
FT 300 300 N -> D (IN ALLELE B).
FT 301 301 M -> I (IN ALLELE B).
FT 301 301 L -> S (IN REF. 2 AND 3).
FT 25 25 S -> P (IN REF. 2 AND 3).
FT 36 36 I -> T (IN REF. 2 AND 3).
FT 239 239
SQ SEQUENCE 336 AA; 36658 MW; 7D879662607C356E CRC64;

Query Match 28.9%; Score 781.5; DB 1; Length 336;
Beet Local Similarity 49.8%; Pred. No. 5e-43; Indels 39; Gaps 9;
Matches 161; Conservative 50; Mismatches 73;

133 LLOG---OSLTLTLESPPGSS-----PSVQCRSPRGKNIQSGKTLVSQLELDGSGTW-- 182
Db LVKGFPESVTVTMNSGSLSSSVHFFPAL-----LQSG-LVTMSSSVTVPSSTWPS 77
Qy 133 ---TCTVLD--NQKVFEEKID-----IVPCPAPEPKSCDKTHTC--PELLGSPVFLF 228
Db 78 QTVTVSVAPASSTVYDKLSPSGPISITNCP-----PCKECHCKCPAPNLEGSPVFI 132
Qy 229 PPKPDTLMIISTPEVTGVVVDVSHEDPEVKENWYVDGVEVNAATKREBDYNTYRV 288

```

```

Db 133 PPKPDTLMIISTPEVTGVVVDVSHEDPEVKENWYVDGVEVNAATKREBDYNTYRV 192
Qy 289 SVLTVLHDDMLNGEKYKKVSNKALPAPIEKTISAKQREPPQYTLPPSRDELTKQV 348
Db 193 STLPIDQDMNSGKEFKCKVNNKDLPSPIERTISKIKLVAPQVYTLPPAEQLSRKV 252
Qy 349 SVTCLVKGFPSPDI AVENESNGQPENNYKTPPVLDSDGSPFLYSKLTVDKSRMQGVF 408
Db 253 SVTCLVGFNPDLSVETTSNGHTBENKDYAPVLDSDGSFTISKLNKTKSKWETDSF 312
Qy 409 SCVWHEALHNHYTKSLSLSPG 431
Db 313 SCVWHEALHNHYTKSLSLSPG 335

RESULT 29
CD4_SAISC STANDARD; PRT; 457 AA.
ID CD4_SAISC
AC 029037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN CD4.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxId=9521;
RN [1]
RP SEQUENCE FROM N.A.
RA Tachumi M., Hashimoto O.;
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D86588; BA013131.1; -.
DR HSBP, P01730; IMR.
DR GO: GO:0042101; C: T-cell receptor complex; ISS.
DR GO: GO:0015026; F: coreceptor activity; ISS.
DR GO: GO:0042289; F: MHC class II protein binding; ISS.
DR GO: GO:0006955; P: immune response; ISS.
DR GO: GO:0045086; P: positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P: T-cell differentiation; ISS.
DR GO: GO:0045058; P: T-cell selection; ISS.
DR GO: GO:0007159; P: transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003596; IG_V.
DR Pfam: PF00047; Ig; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IG; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KM Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 457 BY SIMILARITY.
FT DOMAIN 26 395 T-CELL SURFACE GLYCOPROTEIN CD4.
FT TRANSMEM 396 417 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).

```



```

FT DOMAIN 26 125 IG-LIKE V-TYPE.
FT DOMAIN 126 202 IG-LIKE C2-TYPE 1.
FT DOMAIN 123 316 IG-LIKE C2-TYPE 2.
FT DOMAIN 317 373 IG-LIKE C2-TYPE 3.
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 327 369 BY SIMILARITY.
FT LIPID 418 418 S-palmitoyl cysteine (By similarity).
FT LIPID 421 421 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 457 AA; 50871 MW; 57EED6344005A015 CRC64;

```

```

Query Match 27.5%; Score 743; DB 1; Length 457;
Best Local Similarity 40.1%; Pred. No. 2.1e-40;
Matches 209; Conservative 56; Mismatches 150; Indels 106; Gaps 20;

```

```

OY 1 MNRGVPRHLLVLTQALLPAATQGNKVLKGGDTVELTCTASQKSIQFMKNSQIK 60
DB 1 MNGGIPFRLHLLVLTQALLPAVTHGKTVLKGKGEVVELPCETSLKKNVPEHMKTSQIK 60
OY 61 ILNGSGSLTKGSPSLNDRASRSLSMNGKFPILIKLTKEDSDTYICEVEDQKERYOL 120
DB 61 ILGVQNVFVTRGQSLTRIDSKSSWGRSFPILIKARLESDTYICEVESKKEVEL 120
OY 121 LVFGLTANSDTHLQGGSLTTLTLESPGSSPVOCSPRGNIOGKTLVSQLELDPSG 180
DB 121 QVFGLTANSDTHLQGGSLTTLTLESPGSSPVOCSPRGNIOGKTLVSQLELDPSG 180
OY 181 TWCTCTVLOKKVEFKIDIVCPAPEPKSCNHTCEPILGGSFVLEFP-PRKDTLMIS 239
DB 181 TWCTCTVLOKKVEFKIDIVCPAPEPKSCNHTCEPILGGSFVLEFP-PRKDTLMIS 239
OY 181 TMKCTVFOHLELV-FEINIVLVAQOASS-----TVYKKEGEVDFPFLAFAETLTGS 234
DB 181 TMKCTVFOHLELV-FEINIVLVAQOASS-----TVYKKEGEVDFPFLAFAETLTGS 234
OY 240 RTPETVCVVVDVSHDEPEVKMNVVDGVEYNATKREPEEYNTYRVSVLTYLHADM 299
DB 240 RTPETVCVVVDVSHDEPEVKMNVVDGVEYNATKREPEEYNTYRVSVLTYLHADM 299
OY 235 G-----ELCW-----QARASSSSKSWITFNLTKEVYKLT----- 266
DB 235 G-----ELCW-----QARASSSSKSWITFNLTKEVYKLT----- 266
OY 300 NGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGF 359
DB 300 NGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGF 359
OY 267 --QPKKLMGKK--LPHLTLAQLPQYASSGNFTL-----ALKG--- 302
DB 267 --QPKKLMGKK--LPHLTLAQLPQYASSGNFTL-----ALKG--- 302
OY 360 SDIAVESWESNQEPENNYTTPPVLDSDGSFFLYSKLTVDKSRMGOGVNFCGVNHEALHN 419
DB 360 SDIAVESWESNQEPENNYTTPPVLDSDGSFFLYSKLTVDKSRMGOGVNFCGVNHEALHN 419
OY 303 -----KT-----GKLHGEVNLVYMRVTQONNL-TCEVWGP----- 332
DB 303 -----KT-----GKLHGEVNLVYMRVTQONNL-TCEVWGP----- 332
OY 420 HYTKSLSLSPGLQUDERCARAGDELDGWTTP-PRASALPAPPGSALPDPQTLSAL 478
DB 420 HYTKSLSLSPGLQUDERCARAGDELDGWTTP-PRASALPAPPGSALPDPQTLSAL 478
OY 333 --TSPKMLLS--LKLNEQAKVSKRE-KAVVNLNPEPGAWOCLLSDSGQVLLSK-PEAL 386
DB 333 --TSPKMLLS--LKLNEQAKVSKRE-KAVVNLNPEPGAWOCLLSDSGQVLLSK-PEAL 386
OY 479 P-DPPAASALPALAVISFLGL-----GLGV-ACVLAATR 512
DB 479 P-DPPAASALPALAVISFLGL-----GLGV-ACVLAATR 512
OY 387 PTRSPVQ--PWLVIYGVAGLLAFTGLGIFLCVRCRHR 424
DB 387 PTRSPVQ--PWLVIYGVAGLLAFTGLGIFLCVRCRHR 424

```

```

RESULT 30
CD4_CANFA STANDARD; PRT; 463 AA.
AC P33705;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell) surface antigen
DE T4/Leu-3).
GN CD4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 13-463 FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
MEDLINE=93192324; PubMed=7916632;

```

```

RA Milde K.F., Conner G.E., Minz D.H., Alejandro R.;
RT "Primary structure of the canine CD4 antigen.";
RL Biochim. Biophys. Acta 1172:315-316(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
RX MEDLINE=94378217; PubMed=8091416;
RA Gorman S.D., Frewin M.R., Cobbold S.P., Waldmann H.;
RT "Isolation and expression of cDNA encoding the canine CD4 and CD8
RT alpha antigens.";
RL Tissue Antigens 43:184-186(1994).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in macrophages and a subset of
CC T lymphocytes.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, the European
CC Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L06130; AB02295.1; -.
DR EMBL, X68565; -. NOT_ANNOTATED_CDS.
DR HSSP, P01730; IWR.
DR GO, GO:0042101; C:T-cell receptor complex; ISS.
DR GO, GO:0015026; F:coreceptor activity; ISS.
DR GO, GO:0042289; F:MHC class II protein binding; ISS.
DR GO, GO:0006955; P:immune response; ISS.
DR GO, GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO, GO:0030217; P:T-cell differentiation; ISS.
DR GO, GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro, IPR000973; CD4_TcRg.
DR InterPro, IPR007110; Ig-Like.
DR InterPro, IPR003596; Ig_V.
DR Pfam, PF00047; Ig_3.
DR PRINTS, PR00692; CD4TCANTIGEN.
DR SMART, SM00406; IGV; 1.
DR PROSITE, PS50835; IG_LIKE_1.
DR Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 24
FT CHAIN 25 463
FT DOMAIN 25 401
FT TRANSMEM 402 423
FT DOMAIN 424 463
FT DOMAIN 26 124
FT DOMAIN 125 211
FT DOMAIN 212 321
FT DOMAIN 322 378
FT DISULFID 41 109
FT DISULFID 332 374
FT LIPID 424 424
FT LIPID 427 427
FT CARBOHYD 123 123
FT CARBOHYD 168 168
FT CARBOHYD 176 176
FT CARBOHYD 324 324
FT CARBOHYD 329 329
FT CARBOHYD 389 389
SQ SEQUENCE 463 AA; 51639 MW; 95805170CB44833 CRC64;

```

```

Query Match 22.9%; Score 617.5; DB 1; Length 463;
Best Local Similarity 33.5%; Pred. No. 2.2e-32;
Matches 177; Conservative 68; Mismatches 155; Indels 129; Gaps 19;

```

```

QY 1 MNRGVPFRHLLVLTALPPAATGKNVJLGGKDDVETLCTASQKSIQIHFMMKNSNOIK 60
DB 1 MNQGAFAFRLLMLQVMLPAVTPREVREVLGKAGDAVELPQCTSGKKNIHFMVRDSSNVQ 60
QY 61 ILNGSGSPITKPSKLNDRADRSRLMDQGNFPLIIKLIKIDSPITYICEVEDQKEEVL 120
DB 61 ILNGSGSPITKPSKLNDRADRSRLMDQGNFPLIIKLIKIDSPITYICEVEDQKEEVL 119
QY 121 LVEGTLA-----NSDTHLLOGQSILTLTLESPGSSPVQCRSPRGKNIQGGKTLV 171
DB 120 LVEFNLTAKKDSGSSSSSSNIRLLQGGQLTLTLENSSGSPSYQMGPKNSKHGGQNLSTL 179
QY 172 SOLLELDGSGTWTCTVLONOQKVEFKIDIVPCAPAPKSCDKHTCTPELIGSPVLPFPK 231
DB 180 SMPBELQDGGTWTCTIISQSKTEVFNNVLVLAFA---QKSNFTYARE---GDQVESFP-- 232
QY 232 PNDTLMISRTPEVTCVVVDVSHEDPEV--KFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 289
DB 233 -----LSFEDENLVGELRMQAGAS-----SS 254
QY 290 VITVLHODMLNGKEYKCKVSNKALPAPIEKTSKAKQPRE--PQVYTLPSRDELTKNQ 347
DB 295 LL-----WISFTLENRKLSMKEAHAPL-----KLQMKESLPFRFTLPVLSRYAGSG 301
QY 348 VSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPYLDSGSPFLYSKLTVDKSRMQGNV 407
DB 302 ILTLNLALGTLVQEV-----NLYVWRANSQNNL 330
QY 408 FSCSVMEALHNHYQKSLISLPGQLDDETCAEADGDELGMWTDPPASALPAPPTGS 467
DB 331 -TCEVLGP-----TSPLETLS--LNLKEQAQKVK--QOKLYWVVDPEGGT-----WQC 374
QY 468 ALPDPQ---TASAL--PPPPASALPALA-----VISITLGLGIGVAC 506
DB 375 LLSDDKXVLLASLNVSSSPVVIKSWPKFLATLGLLIGLIGLCVPC 423

```

```

RESULT 31
CD4_RABIT
ID CD4_RABIT STANDARD; PRT; 459 AA.
AC P46630;
DT 01-NOV-1995 (rel. 32, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN CD4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92390370; PubMed=151821;
RA Haque B.F., Sawadikosol S., Brown T.J., Lee K., Recker D.P.,
RT "CD4 and its role in infection of rabbit cell lines by human
immunodeficiency virus type 1."
RL Proc. Natl. Acad. Sci. U.S.A. 89:7963-7967(1992).
CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC - SUBUNIT: Associates with p56-lck (By similarity).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: M92840; AAA31198.1; -.
DR PIR: A46254; A46254.
DR HSPB, P01730, 1MR.
DR GO: GO:0042101; C: T-cell receptor complex, ISS.
DR GO: GO:0015026; F: coreceptor activity, ISS.
DR GO: GO:0042289; F: MHC class II protein binding, ISS.
DR GO: GO:0006955; P: immune response, ISS.
DR GO: GO:0045086; P: positive regulation of interleukin-2 biosyn. . . ISS.
DR GO: GO:0030217; P: T-cell differentiation, ISS.
DR GO: GO:0045058; P: T-cell selection, ISS.
DR GO: GO:0071169; P: transmembrane receptor protein tyrosine kin. . . ISS.
DR InterPro: IPR00973; CD4 TCAG.
DR InterPro: IPR003110; IG-like.
DR InterPro: IPR003596; IG_v.
DR Pfam: PF00047; Ig, 2.
DR SMART; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGv, 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KV Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 459
FT DOMAIN 26 396
FT TRANSMEM 397 419
FT DOMAIN 420 459
FT DOMAIN 26 129
FT DOMAIN 130 208
FT DOMAIN 209 318
FT DOMAIN 319 374
FT CARBOHYD 299 299
FT DISULFID 41 113
FT DISULFID 329 370
FT LIPID 420 420
FT LIPID 423 423
SQ SEQUENCE 459 AA; 50886 MW; B323311CBDA0013D CRC64;

```

```

Query Match 22.3%; Score 601.5; DB 1; Length 459;
Best Local Similarity 41.9%; Pred. No. 2.3e-31;
Matches 153; Conservative 47; Mismatches 98; Indels 67; Gaps 11;
QY 1 MNRGVPFRHLLVLTALPPAATGKNVJLGGKDDVETLCTASQKSIQIHFMMKNSNOIK 60
DB 1 MNKRITPQCLLVLPALPPATGKTVYRGAIVELPQSSQKRNISVNMKHAQYK 60
QY 61 ILNGSG-----SFLTKPSKLNDRADRSRLMDQGNFPLIIKLIKIDSPITYICEVEDQKE 116
DB 61 ILNGSGSSSSSFWLKGNSPLSNRVESKKNMWDQGSFPLVIDLMDMDGTYICEVGDKM 120
QY 117 EVOLLVGLTANSOTHLLOGQSILTLTLESPGSSPVQCRSPRGKNIQGGKTLVLSQLEL 176
DB 121 EVELLVFPLTANPTTRLHQSILTLTLEGPVGSPPVQWSPENKIIETGPTCMPKRLT 180
QY 177 QDSGWTCTV-LONQKVEFKIDIVPCAPAPKSCDKHTCTPELIGSPVLPFPKPD 235
DB 181 QDSGTWCHLSFQONKLELDIKIILVGFPAASA-----TYKKEGEVEFSFP----- 229
QY 236 LMISRTPEVTCVVVDVSHEDPEV--KFNWYVDGVEVHNAKTKPREEQYNSTYRVVS 293
DB 230 -----LNFEDSLSGELMWQDGAS----- 249
QY 294 LHQWLNKKEYKCKVSNKALPAPIEKTSKAKQPREPVYTLPPSRDELTKNQVSLT 351
DB 250 SAQSWSPSLDRDRAVSVQKILP---DKIQWSKGLPLPS---LTLPLALHRYAGSGNLSLT 303
QY 352 CLVVG 356
DB 304 -LDKG 307

```

```

RESULT 32
CD4_RAT
ID CD4_RAT STANDARD; PRT; 457 AA.

```

AC P05540; 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3) (W3/25 antigen).  
 GN CD4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxId=10116;  
 RN (1)  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=87175535; PubMed=3104900;  
 RT Clark S.J., Jeffries W.A., Barclay A.N., Gagnon J., Williams A.F.;  
 RT "Peptide and nucleotide sequences of rat CD4 (W3/25) antigen:  
 RT evidence for derivation from a structure with four  
 RT immunoglobulin-related domains.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 84:1649-1653(1987).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 210-393.  
 RX MEDLINE=93262437; PubMed=8493535.  
 RA Brady R.L., Dodson E.J., Dodson G.G., Lange G., Davis S.J.,  
 RA Williams A.F., Barclay A.N.;  
 RT "Crystal structure of domains 3 and 4 of rat CD4: relation to the  
 RT NH2-terminal domain.";  
 RT Science 260:979-983(1993).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M15768; AAA0901.1; -.  
 DR PIR; A27449; A27449.  
 DR PDB; 1CID; 15-JUN-93.  
 DR GLYCOSITE; P05540; -.  
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR000973; CD4-ITAG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; Ig; 2;  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00409; IG; 2;  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;  
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate; 3D-structure.  
 FT SIGNAL 1 27  
 FT CHAIN 28 457  
 FT DOMAIN 28 394  
 FT TRANSMEM 395 417  
 FT DOMAIN 418 457  
 FT DOMAIN 128 127  
 FT DOMAIN 128 206  
 FT DOMAIN 207 316  
 FT DOMAIN 317 374  
 FT CARBOHYD 186 186  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 T-CELL SURFACE GLYCOPROTEIN CD4.  
 EXTRACELLULAR (POTENTIAL).  
 CYTOPLASMIC (POTENTIAL).  
 IG-LIKE V-TYPE.  
 IG-LIKE C2-TYPE 1.  
 IG-LIKE C2-TYPE 2.  
 IG-LIKE C2-TYPE 3.  
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DISULFID 43 111 BY SIMILARITY.  
 FT DISULFID 158 187 BY SIMILARITY.  
 FT DISULFID 328 370 BY SIMILARITY.  
 FT LIPID 418 418 S-palmitoyl cysteine (By similarity).  
 FT LIPID 421 421 S-palmitoyl cysteine (By similarity).  
 FT STRAND 213 217  
 FT TURN 218 219  
 FT STRAND 222 225  
 FT STRAND 235 243  
 FT STRAND 252 258  
 FT TURN 259 260  
 FT STRAND 261 265  
 FT STRAND 274 275  
 FT STRAND 278 278  
 FT TURN 279 280  
 FT STRAND 282 285  
 FT HELIX 290 292  
 FT STRAND 294 301  
 FT STRAND 306 319  
 FT STRAND 325 331  
 FT STRAND 338 344  
 FT TURN 345 346  
 FT STRAND 349 353  
 FT STRAND 357 361  
 FT STRAND 367 374  
 FT TURN 375 376  
 FT STRAND 377 385  
 SQ SEQUENCE 457 AA; 51437 MW; 477BE157D30954C1 CRC64;  
 Query Match 18.3%; Score 495; DB 1; Length 457;  
 Best Local Similarity 32.8%; Pred. No. 1,5e-24;  
 Matches 137; Conservative 57; Mismatches 132; Indels 92; Gaps 14;  
 QY 1 MARGVPRRH--LLVQLALLPAAAGNKKVYLGGKGVLTCTTASQKSIQPHKNSNQ 58  
 DB 1 MCRGFSFRHLLPILLQLSKLVLTGQKTVLGEKGSALPEBSTSRASFAWKSSDQ 60  
 QY 59 IKILNGSGSLTGPSTKLNDRADSRSLMDQGNFPLIKLKIEDSDTYCEVEDQEEV 118  
 DB 61 KTLGVNKKLLIGSLSELYSPDSRKNAWEGSPILLINKLRMDSQTYVCLENKKEEV 120  
 QY 119 QLVFGLTANSDPHLLQGOSLTTLLES--PGSSPSVOCSPRGNIOGGKTLISVSOLEQ 177  
 DB 121 ELWVFRFTFNRGRLLQGSQSLTILDSNPKVSPPIECCKKSNIVDSKAFTHSLRIQ 180  
 QY 178 DSGTWCTVLQNGKVEF--KIDIVPCAPAPKSCDKTHTCPBLGGPSVFLPPPKDPT 235  
 DB 181 DSGIWCCTVLNOKKISFDKLSVL-----GFASTSIYAKSEGSARFSEFP----- 227  
 QY 236 LMTSRPEVTCVVVDVSHEDPEVKFMWYVNDGVEVHNAKTKPREQQNNTYRVISVLTVA 235  
 DB 228 -----LNIGESLQGELEW-----KAEKAPSS----- 249  
 QY 296 QDWLNGKEYKCKVSNALPAPIEKTIISKAKGQPREPOVYTLPPSRDELTRKNQVSLTCLVK 355  
 DB 250 QSWITTSIKKQKYS-----VQKSTSNPKFQJSE-----TLF-----LTIQI- 285  
 QY 356 GEPYSDIAVESNGPENNYKTPPYLDSGSPFLYSKLTVDKSRMQGNVFSQVM 413  
 DB 286 ----PQSIQFASG-----NLTLT---LDR-GILYQEVNLVWKKVQPPDSNTLTCEVM 331  
 RESULT 33  
 CD4\_MOUSE STANDARD; PRT; 457 AA.  
 ID-CD4\_MOUSE  
 AC P06332;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3) (T-cell differentiation antigen L3T4).

GN CD4.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87018845; PubMed=3094146;  
 RA Tourville B., Gorman S.D., Field E.H., Hunkapiller T., Parnes J.R.;  
 RT "Isolation and sequence of L3T4 complementary DNA clones: expression  
 in T cells and brain.";  
 RL Science 234:610-614(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87115821; PubMed=3027575;  
 RA Littman D.R., Gettner S.N.;  
 RT "Unusual insertion in the immunoglobulin domain of the newly isolated  
 murine CD4 (L3T4) gene.";  
 RL Nature 325:453-455(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88152875; PubMed=3326818;  
 RA Parnes J.R., Hunkapiller T.;  
 RT "L3T4 and the immunoglobulin gene superfamily: new relationships  
 between the immune system and the nervous system.";  
 RL Immunol. Rev. 100:109-127(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=88041159; PubMed=2823269;  
 RA Gorman S.D., Tourville B., Parnes J.R.;  
 RT "Structure of the mouse gene encoding CD4 and an unusual transcript  
 in brain.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7644-7648(1987).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98112780; PubMed=9445485;  
 RA Anselmi-Lati M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,  
 Lu J., Gorrell J.H., Chinnault A.C., Belmont J.W., Miller W.,  
 RA Gibbs R.A.;  
 RT "Comparative sequence analysis of a gene-rich cluster at human  
 chromosome 12p13 and its syntenic region in mouse chromosome 6.";  
 RL Genome Res. 8:129-401(1998).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RX MEDLINE=23388257; PubMed=12477932;  
 RA Streuberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Kleuener R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buecaw K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshimaki S., Carinci P., Prange C.,  
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Boeak S.A., McEwen P.J., McKernan K.O., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [7]  
 RP SEQUENCE OF 27-43.  
 RX MEDLINE=8616694; PubMed=3082751;  
 RA Claesson B.J., Tesgaratos J., Kirsbaum L., Maddox J., McKay C.R.,  
 RA Brandon M., McKenzie I.F.C., Walker I.D.;

RT "The L3T4 antigen in mouse and the sheep equivalent are  
 RT immunoglobulin-like.";  
 RL Immunogenetics 23:129-132(1986).  
 RN [8]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=86233454; PubMed=3086886;  
 RA Claesson B.J., Tesgaratos J., McKenzie I.F.C., Walker I.D.;  
 RT "Partial primary structure of the T4 antigens of mouse and sheep:  
 RT assignment of intrachain disulfide bonds.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4499-4503(1986).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=P06332-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Brain-specific;  
 CC IsoId=P06332-2; Sequence=VSP\_002489;  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; M36850; AAA39401.1; -;  
 DR EMBL; M3816; AAA37267.1; -;  
 DR EMBL; X0483; CAA28539.1; -;  
 DR EMBL; M36851; AAA39402.1; -;  
 DR EMBL; M17080; AAA37403.1; -;  
 DR EMBL; M17078; AAA37403.1; JOINED.  
 DR EMBL; M17079; AAA37403.1; JOINED.  
 DR EMBL; AC002397; AAC36010.1; -;  
 DR EMBL; BC039137; AAC39137.1; -;  
 DR PIR; A02110; RMMST4.  
 DR HSSP; P01730; IMBR.  
 DR MGD; MGI:88335; Cd4.  
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;  
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate;  
 KW Alternative splicing.  
 FT SIGNAL 1 26  
 FT CHAIN 27 457  
 FT DOMAIN 27 394  
 FT TRANSMEM 395 417  
 FT DOMAIN 418 457  
 FT DOMAIN 27 128  
 FT DOMAIN 129 207  
 FT DOMAIN 208 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 187 374  
 FT CARBOHYD 298 298  
 FT CARBOHYD 323 323  
 FT N-TERMINAL 1 26  
 FT T-CELL SURFACE GLYCOPROTEIN CD4.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT IG-LIKE V-TYPE.  
 FT IG-LIKE C2-TYPE 1.  
 FT IG-LIKE C2-TYPE 2.  
 FT IG-LIKE C2-TYPE 3.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 42 112
FT DISULFID 159 188
FT DISULFID 328 370
FT DISULFID 418 418
FT LIPID 421 421 S-palmitoyl cysteine (By similarity).
FT VARSPLIC 1 240 S-palmitoyl cysteine (By similarity).
FT VARSPLIC 1 240 Missing (in isoform 2).
SQ SEQUENCE 457 AA; 51296 MM; 181DAV527CB00F33 CRC64;

Query Match 17.6%; Score 475; DB 1; Length 457;
Best Local Similarity 52.8%; Pred. No. 2.8e-23;
Matches 104; Conservative 32; Mismatches 55; Indels 6; Gaps 5;

QY 1 MNRGVPRFH-DLLVLTALLPATQGNKVLGKGDVTELTCTASOKKSIFPHKNSNQI 59
DB 1 MCRALSLRRLDLLLQLSLAVQKTLVKGESSELPFCSSQKKITVFTKFSQDQ 60
QY 60 KILGNQG-SPLTKG--PSKLNDRADSRSLMDQGNPILINKLIEDSDTYICEVEDQKE 116
DB 61 KILGNHKGKVLIRGSPSQF-DREDSKKGAWKESFPLINKLIMEDSQTYICELENRKE 119
QY 117 EVOLVRELTRANSPTHLQSGSLTTLLES-PPGSSPSQGRKPRGKNIQGGKTLVSQLE 175
DB 120 EVELWVFKVTPSPGTSLSLQSGSLTTLTDSNSKVNPNLTECHKKGVSSKVLMSNLR 179
QY 176 LQDSGTWCTVTLQNOQK 192
DB 180 VQSDSPMNCVTTLQOKK 196

```

RESULT 34  
EPC\_MOUSE STANDARD; PRT; 421 AA.

```

ID EPC_MOUSE STANDARD; PRT; 421 AA.
AC P06336; P01856;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 15-epi chain C region.
DB Ig epsilon chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84236092; PubMed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RT comparison with the human epsilon gene sequence.";
RL EMBO J. 1:1117-1123(1982).
RN [2]
RP REVISIONS.
RA Honjo T.;
RL Submitted (APR-1986) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE=8311774; PubMed=6818553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
RT chain cDNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X01857; CA25977.1; -
DR EMBL; X01857; CA25978.1; -
DR PIR; A02144; EHMS.

```

```

DR PIR; A02144; EHMS.
DR HSP; P01854; 1IGE.
DR InterPro: IPR007110; IG_1like.
DR InterPro: IPR003597; IG_c1.
DR InterPro: IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; glycoprotein.
FT NON TER 1 1
FT DOMAIN 1 90 CH1.
FT DOMAIN 91 197 CH2.
FT DOMAIN 198 304 CH3.
FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MM; 8F909E1F30A06B47 CRC64;

```

Query Match 14.4%; Score 388; DB 1; Length 421;  
Best Local Similarity 28.3%; Pred. No. 9e-18;  
Matches 119; Conservative 77; Mismatches 138; Indels 86; Gaps 20;

```

QY 38 ELCTASQKKSIFPHKNSQIKILGNQSGFLTKGSKLNDRADSRSLMDQGNPILITK 97
DB 55 ELKVTTSQVTS-----WGSKAK-----NFTCVTHPSFNSRT-----ILVR 92
QY 98 NLKIEDSDTYI-----CEVEDQKEVOL--LVFGILTAN--SDTHLQSGSLTLESPPGS 149
DB 93 PVNITEPTELEHSSCDPNAFHTIQYCRTHILNDVSVMLMDREITDYL----- 146
QY 150 SPVSQGRSPRGKNIQGGKTLVSQLELQDSGTWCTVTLQNOQKVEPKIDIVCPAPPKS 209
DB 147 AQTVLILKE-EGKLASCSKLNITEQOMWSESTFTCKV--TSQGVVDLAHRRCPDHEPR- 202
QY 210 CDKTHCPPELLGSPSYFLPPPKKDTLMISRTDEVTCVVDV-SHEPPEYKEN----- 261
DB 203 -----GVITLYLPPSLD-LYQGAPEKLTCLVVDLESKKNVNTMNOEKKTSV 249
QY 262 ----WYVDGEVNAKTKPREEOYNSTYRVSVTLVLDPMNGKEYCKCVSKALPAP1 317
DB 250 SASQWY---TKHN-----NATTSITSLPVAADMIGYQCLVHPDPFKPI 236
QY 318 EKTISRAKQPREPOVYTLPPSRDELTKQVSLTCLVKGFPSPDAVAVESNQEPENN- 375
DB 297 VRSITTPGGRSAPVAVYVPPPREE-SEDKRITLCLIQNFPPDISVQMGDKGLSNQ 355
QY 376 YKTTTPVLSDSG---FLYSKTLTVDSKRNQGNVSCSYMEHALN-HTQKSLSLSPG 421
DB 356 HSTTTP-LKSNNGNQGFIFSRLEVAKTMTQKQFCOVIHIALOKPRKLEXTISTSLG 414

```

RESULT 35  
MOC\_HUMAN STANDARD; PRT; 454 AA.  
ID MOC\_HUMAN  
AC P01871;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ig mu chain C region.  
OS Homo sapiens (Human).  
GN IGHM.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 1-434 FROM N.A.  
 RX MEDLINE=90332450; PubMed=2115966;  
 RA Friedlander R.M., Nussenzweig M.C., Leder P.;  
 RT "Complete nucleotide sequence of the membrane form of the human Igm heavy chain.";  
 RT Nucleic Acids Res. 18:4278-4278(1990).  
 RN [2]  
 RP SEQUENCE (WALDENSTROM'S MACROGLOBULIN GAL).  
 RX MEDLINE=75059123; PubMed=4803843;  
 RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;  
 RT "The primary structure of a monoclonal Igm-immunoglobulin (macroglobulin Gal.). II: the amino acid sequence of the H-chain (mu-type), subgroup H III. Architecture of the complete Igm-molecule.";  
 RT Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).  
 RN [3]  
 RP REVISIONS (GAL).  
 RX MEDLINE=81066716; PubMed=6777162;  
 RA Mhaesco E., Barnikol-Watanabe S., Barnikol H.U., Mhaesco C., Hilschmann N.;  
 RT "The primary structure of the constant part of mu-chain-disease protein BOT.";  
 RT Eur. J. Biochem. 111:275-286(1980).  
 RN [4]  
 RP SEQUENCE (WALDENSTROM'S OU), DISULFIDE BONDS, AND CARBOHYDRATES.  
 RX MEDLINE=74005511; PubMed=4742735;  
 RA Putnam F.W., Florent G., Paul C., Shinoda T., Shimizu A.;  
 RT "Complete amino acid sequence of the mu heavy chain of a human Igm immunoglobulin.";  
 RT Science 182:287-291(1973).  
 RN [5]  
 RP PARTIAL SEQUENCE FROM N.A.  
 RX MEDLINE=82059479; PubMed=6795593;  
 RA Rabibite T.H., Forster A., Milestein C.P.;  
 RT "Human immunoglobulin heavy chain genes: evolutionary comparisons of C mu, C delta and C gamma genes and associated switch sequences.";  
 RT Nucleic Acids Res. 9:4509-4524(1981).  
 RN [6]  
 RP SEQUENCE OF 299-387 AND 438-454 FROM N.A.  
 RX MEDLINE=81077306; PubMed=6777778;  
 RA Dolby T.W., Devuono J., Croce C.M.;  
 RT "Cloning and partial nucleotide sequence of human immunoglobulin mu chain cDNA from B cells and mouse-human hybridomas.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 77:6027-6031(1980).  
 CC -1- MISCELLANEOUS: All 4 combinations of the S/G and V/G polymorphisms at positions 192 and 216 have been observed in human mu chains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, X17115; CAA34971.1; ALT SEQ.  
 CC EMBL, X57086; -, NOT\_ANNOTATED\_CDS.  
 CC HSSP, P01857, 1FC1.  
 CC Genew; HGNC:5541; IGHM.  
 CC MIM, 147020, -.  
 CC GlycosuiteDB; P01871, -.  
 CC GO; GO:0005624; C:membrane fraction; NAS.  
 CC GO; GO:0003823; P:antigen binding; TAS.  
 CC GO; GO:0006955; P:immune response; NAS.  
 CC InterPro; IPR003587; Ig\_L1.  
 CC InterPro; IPR007110; Ig\_L1like.  
 CC InterPro; IPR003587; Ig\_L1.  
 CC Pfam; PF00047; Ig\_4.  
 CC SMART; SM00407; IGH1, 3.  
 CC PROSITE; PS00835; IG\_LIKE; 4.

DR PROSITE; PS00290; IG\_MHC; 3.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Polymorphism.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 105 CH1.  
 FT DOMAIN 106 218 CH2.  
 FT DOMAIN 219 324 CH3.  
 FT DOMAIN 325 454 CH4.  
 FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN).  
 FT DISULFID 28 88  
 FT DISULFID 135 198  
 FT DISULFID 215 215 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 245 304 INTERCHAIN (WITH HEAVY CHAIN IN ANOTHER SUBUNIT).  
 FT DISULFID 292 292  
 FT DISULFID 352 414 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 453 453 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 46 210 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 441 441 /FTid=CAR\_000219.  
 FT VARIANT 192 192 S -> G.  
 FT VARIANT 192 192 /FTid=VAR\_003903.  
 FT VARIANT 216 216 V -> G (in dbSNP:12365).  
 FT VARIANT 216 216 /FTid=VAR\_003904.  
 FT VARIANT 216 216 /FTid=VAR\_003904.  
 FT VARIANT 216 216 /FTid=VAR\_003904.  
 SQ SEQUENCE 454 AA; 49556 MW; 21EC72BAD56922E CRC64;  
 Query Match 14.2%; Score 384.5; DB 1; Length 454;  
 Best Local Similarity 24.9%; Pred. No. 1.7e-17;  
 Matches 114; Conservative 73; Mismatches 164; Indels 107; Gaps 15;  
 QY 36 TWELTCTASQ--KSIQTHKNSNOIKILGNQSTLTGPKSLNDRASRLMDQGNP 93  
 DB 23 SVAAGCLADQFLPDSITFSWKYKNNSDISSTRG-----FP 57  
 QY 94 LIHKLIKEDSDTYCEVED-----QKEEYQLVFLGLTANSDDHLLQGSLTLLESPPGS 149  
 DB 58 SVLRGKTAANSQVILPSPKDMQGTDEHVCKVQHPNGKKNV---PLPIALHPKV 113  
 QY 150 SPVSOCR-----SPR-----GKNIQGGKTLVSQLELDQSG 180  
 DB 114 SYFVPRPGFFGNPRSKSKLICQATGSPRQVSWLABGKQVSGVTTDQQAIAKESG 173  
 QY 181 -----TWCTVLQNRKVEFKIDIVCPAPEPKSCDKHTCPKL 219  
 DB 174 PTTVKTSTLTATKESDMLSQSWFTCRV--DHRGLTFQONASMCVDPDPTAIR----- 224  
 QY 220 LGPSVFLPFPKPKDTLMISRPBYTCVVNVSHDEPKFPMVYDVGVNNAKTKPRE 279  
 DB 225 ----VFALPES-FASIFLTSTKLTCLVTLDTTYD-SVTLISWTRQNGAIVKTHNISES 277  
 QY 280 QYNSTRVSVYTLVADHMLNKEKCYKSNALPAPIEKTISKAKGP-REPOVYTLPP 338  
 DB 278 HENAFSAVGEISIEDDMNSGERFTCYTHTDLPSPKQTSRKGVLAHRPDVYLLRP 337  
 QY 339 SHDELTL-KNOVSLTGLVKGFPYSDIAVWESNGQP--ENNYKTTTPVID--SDGSFFLYS 393  
 DB 338 AREQNLRESATITCLVTGFSPADVFQVMQGGQGLSPKRYTSAIPMEPQAPGRYFAMS 397  
 QY 394 KLTVDKSRMOGANNVSCVMEHALNNHTQSLSLSPG 431  
 DB 398 ILTVSEEMWTGETYTCVVAHREALPNRTERTVDKSTG 435  
 RESULT 36  
 MISC\_MOUSE  
 ID MISC\_MOUSE STANDARD; PRT; 476 AA.  
 AC P01873;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

OS Ig mu chain C region membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 433-476 FROM N.A.
RX MEDLINE=80222874; PubMed=6771020;
RA Early P., Rogers J., Davis M., Calame K., Bond M., Wall R.,
RA Hood L.;
RT "Two mRNAs can be produced from a single immunoglobulin mu gene by
RT alternative RNA processing pathways.";
RL Cell 20:313-319(1980).
[2]
RP SEQUENCE OF 410-476 FROM N.A. (MELONA MOPC 104E).
RX MEDLINE=80222873; PubMed=6771019;
RA Rogers J., Early P., Carter C., Calame K., Bond M., Hood L.,
RA Wall R.;
RT "Two mRNAs with different 3' ends encode membrane-bound and secreted
RT forms of immunoglobulin mu chain.";
RL Cell 20:303-312(1980).
-1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=During differentiation, B lymphocytes switch from
CC expression of isoform Membrane-bound to isoform Secreted;
CC Name=Membrane-bound;
CC IsoId=P01873-1; Sequence=Displayed;
CC Name=Secreted;
CC IsoId=P01872-1; Sequence=External;
CC -1- MISCELLANEOUS: The sequence of residues 1-409 is assumed to be
CC identical with the corresponding region of the secreted form.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/).
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; V00821; CAA24202.1; -.
DR PIR; A02167; MHMSM.
DR HSP; P01857; 1PCL.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; TcG1; 2.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON TER 1
FT DOMAIN 1 105 CH1.
FT DOMAIN 106 217 CH2.
FT DOMAIN 218 324 CH3.
FT DOMAIN 325 436 CH4.
FT TRANSMEM 456 473 POTENTIAL.
FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 28 89
FT DISULFID 136 199 BY SIMILARITY.
FT DISULFID 216 216 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 246 305 BY SIMILARITY.
FT DISULFID 293 293 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 353 415 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .).
FT CARBOHYD 281 281 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 476 AA; 52656 MM; 8D476575A5204071 CXC64;

Query Match 13.8%; Score 373.5; DB 1; Length 476;
Best Local Similarity 25.2%; Pred. No. 8.9e-17;

```

Matches	121,	Conservative	87,	Mismatches	157,	Indels	115,	Gaps	23
QY	37	VELTCTLSQ--KSIQTHMKNSNQIKLIGNQSFUTKCPKLNDRADRSRSLMDQNF--	92						
Db	24	VAMGCLARDELPLPSTISFTWYNNTEVYIGIRTEPT-----LRTGGKYLIA	68						
QY	93	---PLIKNLKIEDSDPY-ICEVEDQKEEVLVFGFLTANSDTHL--LOGQSLLTYLE	144						
Db	69	TSQVLLSPKSLIESDELYVKIH-----YG-GKRNDLHPIPAVAKMNNVAVF	117						
QY	145	SPP-----GSSP---SVQCR---SPR-----GKNIOGG-----	166						
Db	118	VPPRDGFGSGAPRPSKSLICEATNPTPKPIVSWMLKDGLVSGFTTDPVTIENKSTPQT	177						
QY	167	----KTISSVQLKSDSGTWCCTYALONOKVFKIDYIPCPAPEPKSCDKTHTCBELGG	222						
Db	178	YKVI STLISIDIMLNLNVTCR--DHRGLTFLKANSSTCAASPT-----DIL--	225						
QY	223	PSVFLFPKPCKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGEVHNAAKTPKEQYN	282						
Db	226	--FTTIPPSFAD-IFLSKSNLITCLVSNLMTAYE--TLNISMSQSGEPLETIKIKIMESHPN	281						
QY	283	STYAVSVLTLVHODMYLNGKEKKYKCVSNKMLPAPLEKTIISAKQAPR-----QVYVLP	337						
Db	282	GTSFAKGVASVCVADMNNRKEFVCTLVTHRDLPSPQKFKISL---PNEVHKNPVAVLLP	337						
QY	338	PSRDELTKNOVSLTCLVKGFPSPDIADVEMESNQ--PENNYKTTTPVLD--SDGSFPLY	392						
Db	338	PARQQLMLRSATVATCLVKGSPADISVQMLQRGQLLPQEKYVTSAPMRPGARGQGFETH	397						
QY	393	SKLVYDSRRQOGVNFSCSWNHGELNNHYQKSLISRGQLDDETCAENQGEIDGLMTT	452						
Db	398	SILVLTVEENMSGTYTCVGHENLPHLVETRYVDSK---TGEEVNAESEGFEFNMTT	452						

```

RESULT 37
EPC_RAT
ID_EPC_RAT          STANDARD;          PRT;          429 AA.
AC      P01855;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Ig epsilon chain C region.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RC      STRAIN=LOU/C/MSL;
RX      MEDLINE=83064537; PubMed=6292865.
RA      Hellman L., Petersson U., Engstroem A., Karlsson T., Bennich H.;
RT      "Structure and evolution of the heavy chain from rat immunoglobulin
RT      E.";
RN      Nucleic Acids Res. 10:6041-6049 (1982).
RP      [2]
RX      SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RA      MEDLINE=83182019; PubMed=6820340;
RX      Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT      "A cloned cDNA probe for rat immunoglobulin epsilon chain:
RT      construction, identification, and DNA sequence.";
RN      DNA 1:335-343 (1982).
RP      [3]
RP      SEQUENCE OF 205-306 FROM N.A.
RA      MEDLINE=82174576; PubMed=6803238;
RA      Hellman L., Petersson U., Bennich H.;
RT      "Characterization and molecular cloning of the mRNA for the heavy
RT      (epsilon) chain of rat immunoglobulin E.";
RL      Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268 (1982).
CC      -1- SIMILARITY: Contains 4 immunoglobulin-like domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -

```



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL J00744; AAA41379.1; ALT\_INIT.  
DR PIR A93442; EHRT.  
DR HSSP; P01854; 1IGF.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003597; IG-cl.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; Ig\_4.  
DR SMART; SM00407; IgC1.1.  
DR PROSITE; PS00835; IG-LIKE.4.  
DR PROSITE; PS00290; IG\_MHC.3.  
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.  
KW NON\_TER 1  
FT DOMAIN 1 89 IG-LIKE 1.  
FT DOMAIN 103 201 IG-LIKE 2.  
FT DOMAIN 205 305 IG-LIKE 3.  
FT DOMAIN 314 414 IG-LIKE 4.  
FT DOMAIN 168 168 R -> N (IN REF. 2).  
FT CONFLICT 308 P -> L (IN REF. 2).  
SQ SEQUENCE 429 AA; 48671 MW; D2970B34EFA72B0 CRC64;

Query Match 13.7%; Score 370.5; DB 1; Length 429;  
Best Local Similarity 31.5%; Pred. No. 1.2e-16;  
Matches 88; Conservative 62; Mismatches 102; Indels 27; Gaps 10;

OY 160 GKNIQGGKLSVLSQLELQDSGTWTCTVLQNOKKVEKIDIVCPAPAPKSCDKHTCPRL 219  
DB 160 GLASTYSLNLTITQOQMSSESTCTKVTSGE--NYMATRRCSDDEPR----- 206  
OY 220 LGSPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTKRREE 279  
DB 207 --GVITLYLPPSLD--LYENGTFKLTCLVLDLSESE--NITVMVRRKKSIGASQGRSTK 262  
OY 280 QVNSTYRVSVLTVLHQDLNKGKEYCKSNKALPPIEKTISKAKGQPREPQVYTLPP 338  
DB 263 HHNATTSITSLIPVDKMDIIEGEGYCCRVDPHPFPRIVRSITKAGKSAPEVYVFLP 322  
OY 339 SRDELTKNOVSLTCLVKGFPYSPDIAVEW--ESNGQENNYKTTPTPLDSDG---FLYS 393  
DB 323 EEE--KDKRTILCLIGNFPEDISIQMLQDSLTILKSGHSTTP--LTKNGSNQRFRTS 379  
OY 394 KLTVDKSRWQGNVFSQVMEALHN--HYTKSLSLSPG 431  
DB 380 RLEVTALMTQTFQKFCRVYHEALRPRLKLERITISKSLG 418

RESULT 38  
MUCM RABIT STANDARD; PRT; 479 AA.  
AC P04221;  
DT 20-MAR-1987 (rel. 04; Created)  
DT 01-NOV-1991 (rel. 20; Last sequence update)  
DT 10-OCT-2003 (rel. 42; Last annotation update)  
DE IG mu chain C region membrane-bound form.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_Taxid=9986;  
RN (1)  
RP SEQUENCE FROM N.A. (A2 ALLOTYPE).  
RX MEDLINE=84088930; PubMed=6418803;  
RA Bernstein K.E., Alexander C.B., Reddy E.P., Mage R.G.;  
RT "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain  
of Vha2 allotype: comparisons with Vha1 and membrane mu sequences.";  
JC Immunol. 133:490-495 (1984).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Comment=During differentiation, B lymphocytes switch from

CC expression of isoform Membrane-bound to isoform Secreted;  
CC Name=Membrane-bound;  
CC Name=P04221-1; Sequence=Displayed;  
CC Name=Secreted;  
CC IsoId=P03988-1; Sequence=External;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL K01357; AAA31293.1; -  
DR PIR; A02165; MHRM.  
DR HSSP; P01842; 7FAB.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003597; IG-cl.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; Ig\_4.  
DR SMART; SM00407; IgC1.2.  
DR PROSITE; PS00835; IG-LIKE.4.  
DR PROSITE; PS00290; IG\_MHC.3.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW Alternative splicing; Transmembrane.  
KW NON\_TER 1  
FT DOMAIN 1 106 CH1.  
FT DOMAIN 107 222 CH2.  
FT DOMAIN 223 327 CH3.  
FT DOMAIN 328 458 CH4.  
FT TRANSMEM 459 476 POTENTIAL.  
FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).  
FT DISULFID 28 90 BY SIMILARITY.  
FT DISULFID 137 200 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).  
FT DISULFID 219 219 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).  
FT DISULFID 249 308 BY SIMILARITY.  
FT DISULFID 296 296 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).  
FT DISULFID 356 418 BY SIMILARITY.  
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 479 AA; 52351 MW; 689C637A47BE19FC CRC64;

Query Match 13.7%; Score 369.5; DB 1; Length 479;  
Best Local Similarity 23.8%; Pred. No. 1.6e-16;  
Matches 123; Conservative 79; Mismatches 153; Indels 161; Gaps 23;

OY 21 AATGNKVVLLGKGGDTVELTCTASQ--KKSIOFHW--KNSNQIKILNGSFLTKGPKSL 76  
DB 17 ALTDGILNVAMG-----CLARDFLPSVTFSSFKNSSEI----- 50  
OY 77 NDRADSRSLMNGQNPFLIKNLK-----IEDDTY--ICEVEDQKEVQL 121  
DB 51 -----SSRTV--RTFPVVKRGDKYMATSQVLVPSKDVLYQTEELYLVCKVQHSNRRDLR 102  
OY 122 VEGLTANSDTHLLQGSLLTLTLESPGSSPSVQCR-----SPR-- 159  
DB 103 V-----SPFVDSLELPNNVSVFPFPDPSGSGSTRKSRLLCAQTGSPKQI 147  
OY 160 -----GKNIQGG-----KTLVSQLELDQSGTW-----TCTV-----L 187  
DB 148 SVSWLRDQGVKSVGLTKPVEAFETKGAGPATPSSISMLTITSSDWLSQSLTCRVDRHGI 207  
OY 188 QNOKYVERKIDIVCPAPAPKSCDKHTHCPPELLGSPSVFLPPKPKDTLMISRTPEVTCV 247  
DB 208 FPDKNVSMSSSCSTTPSP-----GIQVFPIAPSPADT--FLSKSARLLCL 250  
OY 248 VVDVSHEDPEVFNMYVDGVEVHNAKT-----KPREQVNSTYRVSVLTVLHQDLNKG 302

```

Db 251 VTDLITVYG-SLINISW-----ASHNGKALDTHMNTESHSPNATFSAMGEASVCAEDMESGE 304
Qy 303 EYKCVSKKALPAPLEKTIISAKGPRE-POVYTLPPSRDEL-KNQVSLICLVKGFPS 360
Db 305 QPTCVTHADLPFPKHTIISREVAKRPAYVYLPAREQLVLESATVTCVLVGFSPA 364
Qy 361 DIAVESNNGOP--ENNKKTPPVLD--SDGSFFLYSKLTVDKSRMOQGNFSCSVMEBA 416
Db 365 DVFVQMGQRGPLSDKVTTSAPAPEPQAPGLYFTHSTLTYTEEDMNGGEFTTCVVGHEA 424
Qy 417 LHNHYTKQSLSLSPGLQDETCAERQDELDGLMTT 452
Db 425 LPHMVTERTVDKS-----TEGEVGAEEGFEFLMTT 455

RESULT 39
MUCB_CANFA STANDARD; PRT; 450 AA.
ID MUCB_CANFA
AC P01874;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig mu chain C region.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 1-177 (MOO).
RX MEDLINE=80077682; PubMed=117299;
RA McCumber L.J., Capra J.D.;
RT "The complete amino-acid sequence of a canine mu chain.";
RL Mol. Immunol. 16:565-570 (1979).
RN [2]
RP SEQUENCE OF 178-450 (MOO).
RX MEDLINE=78180587; PubMed=653360;
RA Wasserman R.L., Capra J.D.;
RT "Amino acid sequence of the Fc region of a canine immunoglobulin M:
inter-species homology for the Igm class.";
RL Science 200:1159-1161 (1978).
DR PIR; A83131; MHDG.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON TER
SQ SEQUENCE 450 AA; 48895 MW; 9DA60DA9D1012F5D CRC64;

Query Match 13.7%; Score 369; DB 1; Length 450;
Best Local Similarity 24.1%; Pred. No. 1.6e-16;
Matches 119; Conservative 80; Mismatches 139; Indels 156; Gaps 20;

Qy 36 TVELTCTASQ--KSGIOFHMKSNQIKILGNQSTLTGSPKLANDRADSRLNDQGNFP 93
Db 22 TVAMGCLARDPLPGSITFSMKYBLSTAINSTRG-----FP 56
Qy 94 LIKKLK-----IEDSDTYI-CEVE-----DOKEVQLVFGLTANSPTHLL 134
Db 57 SVLRGKRVANSQVLPBVDIIQGTDEHIVCKVHBSBKQKVPVPM----- 105
Qy 135 QGQSITLTLESPPGSSPSVQCR-----SPR-----GKNIOGSKT 168
Db 106 -----LTL--PPEVSGFIPRDAFGBPRKSQILCOAGSPRQVMSLRQKQIESGVT 157
Qy 169 LSVSGLLEQDSG-----TWTCT-----TVLONQKVEFKDIV 200
Db 158 TNEVAZAKZSGPTTYKVTSMLTIOEDAMLSQSVFTCVBHRGLTFQONASM----- 210

```

```

Qy 201 PCPAPPEKSCDKTHTCEPLGSPSEVFLPPPKKDTLMISRTPEVTVVDVSHEDPEYK 260
Db 211 -CTSDQGV-----GISIFITIPS-FASINTNTSAKSLCVLTDLATPD-SVTI 254
Qy 261 NWYDVGEVHNATKPREEQNSTYRVVSLTVLHQDLNGKEYCKVSKALPAPLEKT 320
Db 255 SWTEENGALKTHTNISESHPNGTFPSAMGEATVCEBEMESGEQPTCVTHDLPVSKOT 314
Qy 321 ISXKKG--QPREPOVYTLPPSRDEL-TKNQVSLICLVKGFYPSDIAVESNNGO--PENNY 376
Db 315 ISRPKGVAVHNPVYVLPSPREQLDLRESATLSCLTVGSPDVGVQVQKQGVPPUSY 374
Qy 377 KTTPEVLD--SDGSFFLYSKLTVDKSRMOQGNFSCSVMEBALHNYTKQSLSLSPG--- 431
Db 375 VTSAPMEPQAPGLYFTHSTLTYTEEDMNGGEFTTCVVAHESLPRNTERSVDSGTGKPT 434
Qy 432 -----IQDETCAE 440
Db 435 LYNVSLVSDTAGZ 448

RESULT 40
MUCB_HUMAN STANDARD; PRT; 391 AA.
ID MUCB_HUMAN
AC P04220;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig mu heavy chain disease protein (BOT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=84184186; PubMed=6425189;
RA Barnikol-Watanabe S., Mihaesco E., Mihaesco C., Barnikol H.U.,
RA Hilschmann N.;
RT "The primary structure of mu-chain-disease protein BOT. Peculiar
amino-acid sequence of the N-terminal 42 positions.";
RL Hoppe-Seyler's Z. Physiol. Chem. 365:105-118 (1984).
CC -I- MISCELLANEOUS: This protein has no V region homology or CH1
region.
DR PIR; A02163; MHHUBT.
DR HSSP; P01857; 1FC1.
DR GO; GO:0005624; C:membrane fraction; NAS.
DR GO; GO:0003823; F:antigen binding; TAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 42
FT DOMAIN 43 155
FT DOMAIN 156 261
FT DOMAIN 262 391
SQ SEQUENCE 391 AA; 43057 MW; 9100843AF0CF021A CRC64;

Query Match 13.5%; Score 365.5; DB 1; Length 391;
Best Local Similarity 28.0%; Pred. No. 2.3e-16;
Matches 105; Conservative 65; Mismatches 132; Indels 73; Gaps 15;

Qy 110 EVED-----QKEEVQLVFGLTANSPTHLLTLESPPGSSPSVQCR----- 156
Db 18 EAEDRIITKEEARL-----SGRD--MOUTSQPVIALELPKVSVPVPRRGFGFNPKS 68
Qy 157 -----SPR-----GKNIOGSKTLSVQLELDQSG-----TWTCTVLON--- 189
Db 69 KLICQATGFSPPRLQEVMSLRGKQVGSVTTDEVEBAKESGPTTYKVTSTLTIKESDWL 128

```

```
OY 190 -OKVFEKIDIVPCPAPEKSCDKHTTCPELLGSPS-----VLPPEPKDPTLMISRTP 242
DB 129 GQSMFCRVDHGLTQGNASS-----MCGEPDPAIRVFAIPSP-FASIFLTXT 178
OY 243 EVTCVVVDVSHEDPEVKFNNVYDGEVHNAAKTPREEDQNSTYRVSVLTVTHQDMLNCK 302
DB 179 KLTCLVTLDTLVYD-SVTISMTRODGEAVKHTNINISSEHNATFSAVGEASICEDMDSGE 237
OY 303 EKRCVSNKALPAPLEKITSKAKGCP-REPOVYTLPPSDELT-KQVSLTCLVKGFPYS 360
DB 238 RETCTVHTHDLPSPLKQTTISRKQVALHPDYLLPPEAQINLRSAITCLVYGFSPA 297
OY 361 DVAEWSNGCP--ENNYKTTTPVLVD--SDGSFPLSKLTVDKSRVQGNVPSGSVMHEA 416
DB 298 DVFQVMQMGKQPLSPKTYTSAPEPPOAPGRIFAHISILTVSEEWKNGTETTCVVAHEA 357
OY 417 LHNHYTQKSLSLSPG 431
DB 358 LPNRYTERTVDSKSTG 372

RESULT 41
MUC_MOUSE STANDARD; PRT; 455 AA.
AC 101872;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DB Ig mu chain C region secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090;
RX MEDLINE=81076590; PubMed=6255422;
RA Kawakami T., Takahashi N., Honjo T.;
RT "Complete nucleotide sequence of mouse immunoglobulin mu gene and
RL comparison with other immunoglobulin heavy chain genes.";
RN Nucleic Acids Res. 8:3931-3945(1980).
RX MEDLINE=82051295; PubMed=6795090;
RA Goldberg G.I., Varin E.F., Zrolka A.M., Blattner F.R.;
RT "Sequence of the gene for the constant region of the mu chain of
RL Balb/c mouse immunoglobulin.";
RN Gene 15:33-42(1981).
RX MEDLINE=81165562; PubMed=6260591;
RA Auftray C., Rougeon F.;
RT "Nucleotide sequence of a cloned cDNA corresponding to secreted mu
RL chain of mouse immunoglobulin.";
RN Gene 12:77-86(1980).
RX MEDLINE=83075344; PubMed=6816276;
RA Kehy M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
RT "Complete amino acid sequence of a mouse mu chain: homology among
RL heavy chain constant region domains.";
RN Biochemistry 21:5415-5424(1982).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=During differentiation, B lymphocytes switch from
```

```
CC expression of isoform Membrane-bound to isoform Secreted;
CC Name=Secreted;
CC IsoId=P01872-1; Sequence=Displayed;
CC Name=Membrane-bound;
CC IsoId=P01873-1; Sequence=External;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, V00827; -; NOT_ANNOTATED_CDS.
CC PIR, A02166; MHMS.
CC HSSP, P01857; 1FC1.
CC InterPro; IPR007110; IG-1like.
CC InterPro; IPR003597; IG_C1.
CC InterPro; IPR003006; IG_MHC.
CC Pfam, PF00047; Ig_4.
CC SMART, SM00407; IGC1; 2.
CC PROSITE, PS50835; IG-LIKE; 4.
CC PROSITE, PS00290; IG_MHC; 3.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
CC Alternative splicing.
CC NON_TER 1
CC DOMAIN 1
CC FT 105 CH1.
CC FT 106 CH2.
CC FT 218 CH3.
CC FT 324 CH4.
CC DOMAIN 325
CC FT 437 455 C-TERMINAL SEGMENT OF SECRETED FORM.
CC FT 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
CC FT 28 89 BY SIMILARITY.
CC FT 136 199 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
CC FT 216 216 BY SIMILARITY.
CC FT 246 305 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
CC FT 293 293 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
CC FT 353 415 BY SIMILARITY.
CC FT 454 454 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
CC FT 46 46 N-LINKED (GLCNAC. . .).
CC FT 211 211 N-LINKED (GLCNAC. . .).
CC FT 243 281 N-LINKED (GLCNAC. . .).
CC FT 281 281 N-LINKED (GLCNAC. . .).
CC FT 442 442 N-LINKED (GLCNAC. . .).
CC FT 78 78 S -> N (IN MOPC 104E).
CC FT 101 101 H -> Q (IN MOPC 104E).
CC FT 226 226 T -> N (IN TERC183 AND MOPC 104E).
CC FT 258 258 N -> T (IN MOPC 104E).
CC FT 258 258 N -> S (IN TERC183).
CC FT 368 368 L -> K (IN TERC183 AND MOPC 104E).
CC SQ SEQUENCE 455 AA; 50101 MW; 4CB57CB602F9B51 CRC64;

Query Match 13.4%; Score 362; DB 1; Length 455;
Best Local Similarity 25.3%; Pred. No. 4,6e-16;
Matches 116; Conservative 84; Mismatches 149; Indels 110; Gaps 22;

OY 37 VELTCTASQ-KKSIQFMKNSQIKLGNQSGFLTKPSPKLNDRADRRSLMDQGNF-- 92
DB 24 VAMGLADFLPSTISFTWVQNTVEVIGIRTFPT-----LRTGKXYLA 68
OY 93 ---PLIINKKLEDDTY-ICEVEQKEVQLLVGLTANSPTHL-----LQGSILTLTLE 144
DB 69 TSQVLSLSPKSLIEGDEYLVCKH-----YG-ISKNRDLHVPDVAVMNNVNVF 117
OY 145 SPP-----GSSP-----SVOCR-----SPR-----GKNIOQG----- 166
DB 118 VPRPGFGSPAPRKSKLTICELNFTPKPTYSMLKQGLVBSGFTTDDVTIENKSTPOT 177
OY 167 ---KTLVSOQLLELDSSGTWGTCTVLQONOKVEFKIDIVPCPAPEKSCDKHTTCPELLG 222
DB 178 YKVISTLTISIDWLNLVYTCRV--DHRGLTFLKNVSTCAASPST-----DIL-- 225
```

```

QY 223 PSVFLPPKPKDITMISRTPEVTCVAVDVSHEDPEKENVYDGVAVNAKTPREEOYN 282
DB 226 --TFIPSPSPAD-IFLSKSANLTCVSNLATYE-FLNLSMWSQSEPELETIKIMESHFN 281
QY 283 STYRVSATLTVLHODWLNCKEYKCVSKALPAPEKTISSAKGQPRE-----POVYTLR 337
DB 282 GTFSAKGVASVCVEDMNRKKEVCTVTRHDLPSPKKFKISK-----PNTVHKHPAVVYLLP 337
QY 338 PSRDLT-KNGVSLCLVKGFPSPDIAYEWESNGQ--PENNYKTPPVLD--SDGSFPLY 392
DB 338 PARDELNRBSATVTCVVKGSPADISVOMLORGQLLOEKYVTSAPMPEPQAPGFETH 397
QY 393 SKLTVDKSRWQGVNFSQSVNHEALHNYTOKSLSPG 431
DB 398 SILVTVEEMNSGETYTCVVGHEALPHLVTERTVDKSTG 436

RESULT 42
EPC_HUMAN STANDARD; PRT; 428 AA.
AC P01854;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig epsilon chain C region.
GN IGHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OK NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=83168897; PubMed=6300763;
RA Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,
RA Kikuchi M., Sugino Y., Nishida Y., Honjo T.;
RT "Molecular cloning and nucleotide sequencing of human immunoglobulin
RT epsilon chain cDNA."
RL Nucleic Acids Res. 11:719-726(1983).
RN (2)
RP SEQUENCE FROM N.A., AND VARIANT LEU-359.
RX MEDLINE=83001945; PubMed=6288268;
RA Max E.B., Battey J., Ney R., Kirsch I.R., Leder P.;
RT "Duplication and deletion in the human immunoglobulin epsilon gene."
RL Cell 29:691-699(1982).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=84236029; PubMed=6234164;
RA Flanagan J.G., Rabbitts T.H.;
RT "The sequence of a human immunoglobulin epsilon heavy chain constant
RT region gene, and evidence for three non-allelic genes."
RL EMBO J. 1:655-660(1982).
RN (4)
RP SEQUENCE FROM N.A.
RX MEDLINE=84207910; PubMed=6327276;
RA Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;
RT "Long terminal repeat-like elements flank a human immunoglobulin
RT epsilon pseudogene that lacks introns."
RL EMBO J. 1:1539-1544(1982).
RN (5)
RP PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).
RA Benrich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;
RL (in) Bach M.K. (eds.);
RT Immediate hypersensitivity: modern concepts and developments, pp.1-36,
RL Marcel Dekker, New York (1978).
RN (6)
RP SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.
RX MEDLINE=83065234; PubMed=6815656;
RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,
RA Bell L.O., Gould H.J.;
RT "Cloning and sequence determination of the gene for the human
RT immunoglobulin epsilon chain expressed in a myeloma cell line."
RL Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).
RN [7]

```

```

RP 3D-STRUCTURE MODELING.
RX MEDLINE=87089848; PubMed=3796618;
RA Padlan E.A., Davies D.R.;
RT "A model of the Fc of immunoglobulin E."
RL Mol. Immunol. 23:1063-1075(1986).
CC -1- SIMILARITY: Contains 4 immunoglobulin-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L000022; AAB59424.1; ALT_INT.
DR PIR; A22771; EHHU.
DR PDB; 1IGE; 15-JUL-92.
DR PDB; 1FP5; 30-JAN-02.
DR PDB; 1G84; 16-MAY-01.
DR PDB; 100V; 18-SEP-02.
DR Genew; HGNC:5522; IGHE.
DR MIM; 147180; -.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR Interpro; IPR007110; Ig-like.
DR Interpro; IPR003597; Ig cl.
DR Interpro; IPR003006; Ig_MHC.
DR Pfam; PF00407; Ig; 4.
DR SMART; SM00407; IGc1; 4.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure; Repeat.
FT NON_TER 1
FT DOMAIN 6 103
FT DOMAIN 112 210
FT DOMAIN 214 318
FT DOMAIN 324 423
FT DISULFID 14 14
FT DISULFID 15 105
FT DISULFID 29 85
FT DISULFID 121 121
FT DISULFID 135 193
FT DISULFID 209 209
FT DISULFID 239 299
FT DISULFID 345 405
FT CARBOHYD 21 21
FT CARBOHYD 49 49
FT CARBOHYD 99 99
FT CARBOHYD 146 146
FT CARBOHYD 252 252
FT CARBOHYD 275 275
FT VARIANT 359 359
FT STRAND 110 110
FT STRAND 113 117
FT HELIX 122 124
FT TURN 125 125
FT STRAND 130 141
FT STRAND 146 150
FT STRAND 155 155
FT STRAND 159 161
FT STRAND 165 166
FT TURN 168 169
FT STRAND 172 181
FT HELIX 182 186
FT TURN 187 188
FT STRAND 192 196
FT TURN 198 199
FT TURN 201 202
FT STRAND 205 207
FT STRAND 218 222

```

```

FT TURN 226 229
FT TURN 231 232
FT STRAND 236 242
FT STRAND 252 252
FT STRAND 255 257
FT TURN 258 259
FT STRAND 260 260
FT TURN 265 266
FT STRAND 271 274
FT TURN 275 276
FT STRAND 277 285
FT HELIX 288 293
FT TURN 294 294
FT STRAND 297 302
FT TURN 304 305
FT STRAND 310 314
FT STRAND 322 322
FT STRAND 329 329
FT HELIX 334 337
FT STRAND 340 348
FT STRAND 350 351
FT STRAND 356 360
FT TURN 366 369
FT STRAND 374 374
FT STRAND 377 378
FT TURN 380 381
FT STRAND 384 385
FT STRAND 387 393
FT HELIX 394 398
FT TURN 399 400
FT STRAND 404 408
FT STRAND 418 421
SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;

Query Match 13.4%; Score 361.5; DB 1; Length 428;
Best Local Similarity 28.9%; Pred. No. 4.5e-16;
Matches 102; Conservative 53; Mismatches 137; Indels 61; Gaps 10;

QY 90 GNPPLIKKLIKEDSDTYICEVEDQKEVQLVFGILTAN--SDTHLLQGQSLTLTLESPP 147
DB 125 GHFPPITQL-----CLVSGYTPGTINITWLEDQVMDVDL----- 160
QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTMTCTVLQNGKVEFKIDIVPCAPRP 207
DB 161 ----STAATQBELASTOSTELTSQKHLSDRTTYCQVYQHTPE----- 203
QY 208 KSCDKHTCPREL-LGSPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDG 266
DB 204 ----DSTKKCADSNPREVSAVLSRSPFD-LFIRKSPITICLVVDLAPSKGTVALTWSRAS 259
QY 267 VEVHNAKTKPREQVNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKISKAG 326
DB 260 GKVNHSTKEKEKQNGTLVTSTLPLVGTRDWEGETYOCRVTHPHLPALMSTTKTSG 319
QY 327 QPREPOVYTL-----PESRDELTKQNVSLTCLVKGYFPPYDIAVEMESN--GQEPENNYKTT 379
DB 320 PRAAPFVYAPAPFEMWGSRDKRT-----LACLQNPMPEDISVOMLHNEVOLPDARHSTT 374
QY 380 PVLVDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHYT--QKSLSLSPG 431
DB 375 QPKTKGSGFFVSRLEVTTRAEMEQKDEFICRAVHEDASQIVQRAVSVNFG 427

RESULT 43
MUC_RABIT STANDARD; PRT; 458 AA.
AC P03988;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1g mu chain C region secreted form.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A. (A2 ALLOTYPED).
RA MEDLINE=8408930; PubMed=6418803;
RX Bernstein K.E., Alexander C.B., Reddy E.P., Mage R.G.;
RT "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain
of Vha2 allotype: comparisons with Vha1 and membrane mu sequences.";
RL J. Immunol. 132:490-495(1984).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=During differentiation, B lymphocytes switch from
CC expression of isoform Membrane-bound to isoform Secreted;
CC Name=Secreted;
CC IsoId=P03988-1; Sequence=Displayed;
CC Name=Membrane-bound;
CC IsoId=P04221-1; Sequence=External;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; K01357; -; NOT_ANNOTATED_CDS.
DR PIR; A02164; MHRB.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternating splicing.
KW NON_TER
FT DOMAIN 1 106 CH1.
FT DOMAIN 107 222 CH2.
FT DOMAIN 223 327 CH3.
FT DOMAIN 328 458 CH4.
FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 28 90 BY SIMILARITY.
FT DISULFID 137 200 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 219 219 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 249 308 BY SIMILARITY.
FT DISULFID 266 296 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 356 418 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 458 AA; 49897 MW; F338D6A3B28E4864 CRC64;

Query Match 13.3%; Score 360; DB 1; Length 458;
Best Local Similarity 23.8%; Pred. No. 6.2e-16;
Matches 118; Conservative 76; Mismatches 145; Indels 156; Gaps 22;

QY 21 AATGKVKVLLGKKGVTELCTASQ--KKSIOFHW--KSNQIKILGNGSFLTKGPSKL 76
DB 17 ALTDGVLVAMG-----CLARDPLSSVTFSSFKNSRI----- 50
QY 77 NDRADSRSLMDQGNFLLIKLK-----IEDSDTY-ICEVEDQKEVQLL 121
DB 51 -----SSRTV---RFPVYVKRGDKYMATSQVLVPSKDVLCQTEBYLVCKVQHSNRRDLR 102
QY 122 VGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCR-----SPR-- 159

```

[illegible]

RESULT 44	MUC_SUNMU	STANDARD;	PRT;	457 AA.
AC	P20768;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig mu chain C region.			
OS	Sinus murinus (House shrew) (Musk shrew).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Insectivora; Soricidae; Crocidurinae; Suncus.			
OX	NCBI_TaxID=9378;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=layer1;			
RX	MEDLINE=69232144; PubMed=2497033;			
RA	Isihiguro H., Ichihara Y., Namikawa T., Nagatsu T., Kurosawa Y.;			
RT	"Nucleotide sequence of Suncus murinus immunoglobulin mu gene and comparison with mouse and human mu genes.";			
RL	FEBS Lett. 247:317-322(1989).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X13920; CAA32113.1; ALT_INT.			
DR	PIR; S03961; S03961.			
DR	HSSP; P01842; 7FAB.			
DR	InterPro; IPR007110; IG-1-like.			
DR	InterPro; IPR003597; IG_C1.			
DR	InterPro; IPR003006; IG_MHC.			
DR	Pfam; PF00047; Ig; 4.			
DR	SMART; SM00407; IGC1; 3.			
DR	PROSITE; PS50835; IG_LIKE; 4.			
DR	PROSITE; PS00290; IG_MHC; 3.			
KW	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.			
FT	NON TER	1	105	CH1.
FT	DOMAIN	1	220	CH2.
FT	DOMAIN	106	220	CH2.
FT	DOMAIN	221	326	CH3.
FT	DOMAIN	327	457	CH4.

FT	DISULFID	13	13	INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT	DISULFID	27	89	BY SIMILARITY.
FT	DISULFID	136	200	BY SIMILARITY.
FT	DISULFID	217	217	INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT	DISULFID	248	307	BY SIMILARITY.
FT	DISULFID	295	295	INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT	DISULFID	355	417	BY SIMILARITY.
FT	DISULFID	456	456	INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT	CARBOHYD	45		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	113	113	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	212	212	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	276	276	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	283	283	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	457 AA;	50074 MW;	56C0806DA4462E9 CMC64;
Query Match	Similarity	13.1%;	Score 354.5;	DB 1; Length 457;
Best Local Similarity		25.5%;	Pred. No. 1.4e-15;	
Matches	122;	Conservative	71;	Mismatches 163; Indels 123; Gaps 19;

Query Match	13.1#	Score 354.5	DB 1.5	Length 457
Best Local Similarity	25.5#	Pred. No. 1.4e-15		
Match 122	Conservative 71	Mismatches 163	Indels 123	Gaps 19
QY	19	LPAATGKKNVVLGKKGBDVEILTTASQ--KKSIOFHKNSNOIKILGKQSGFLTKGPEKL	76	
DB	17	LPEDTQ-----VTLGCLARDFLPRPVFMKFKXSSSI-----	49	
QY	77	NDRADSRSLMDQGNFLILKNL-----IEDSDTYI-CEVDQKEEYOLL	121	
DB	50	-----SSONIT--NPEVEFTGKYMATSOVLLPSTALIOSTDYITCTHGTGGEKEKK	101	
QY	122	VFGLTASDFTLLQGOSLITLLESPPGSPSPVOCSPRGKNIQGGKTLVSQLEIODSG-	180	
DB	102	V-----ELQVTPELRPVNSIFV--PR--NSFGNHRTSQILQASGF	141	
QY	181	--TWCTVLQNKKEVEPKIDVCPAPEKSCKT-----HTCP	217	
DB	142	SPRTIVMSWLQGEPPVOPSLVSTSAVAEPPGSGPTFRVISRLITENEWLSOREFTCQ	201	
QY	218	EL-----LG-----GSAVFLPPRKQDLMISRTEVCVVVDVSHDEEV	258	
DB	202	ALHAGLTFQKNVSSVCMGDDTSTGISVFLRPPTAN--IFLTQSOALCTLVGLATYD--SL	259	
QY	259	KFNMYVDGVEVHNAKTPREBOVNSTYRVVSVLTVLHODMLNGEKYCKVANKLPAPIE	318	
DB	260	DISMRONGEALQTHNVNISSEHPRSTFTAKGHAIVCREEMWSEGEFTCTVGHSDLPBKL	319	
QY	319	KTISKAKGQRE-PQVYTLPPSRDEL-TKQVSLITCLYKGYRPDIADVEMSGOR--EN	374	
DB	320	QSLRPRQVANDPSPVFLPRAEQOLKLRBSASITCLVCKPFSPRDVFQWQHCGQVDPK	379	
QY	375	NYKTTPPVLD--SPGSPFLYSKLVVDKSRMQGVFSCSVNHALHNHYTOKSLSPG	431	
DB	380	HYVTSNTPPEQNGFLYFVHSILTVSEKDMSSGSEFSCVGHLEALPLSVTEKAVDKTSG	438	





```

RL RN J. Immunol. 143:3632-3640(1989) .
[6]
RP INTERACTION WITH SECTM1.
RX MEDLINE=20119303; PubMed=10652336;
RA Lyman S.D., Escobar S., Rousseau A.-N., Armstrong A., Fanslow W.C.;
RT "Identification of CD7 as a cognate of the human K12 (SECTM1) protein.";
RL J. Biol. Chem. 275:3431-3437(2000) .
CC -1- FUNCTION: Not yet known.
CC -1- SUBUNIT: Interacts with SECTM1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD7 entry.
CC WWW=http://www.ncbi.nlm.nih.gov/prow/cd/cd7.htm.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
-----
CC EMBL; X06180; CAA29546.1; -.
DR EMBL; M37271; AAAS1953.1; -.
DR EMBL; BC009293; AAAH09293.1; -.
DR EMBL; BC013297; AAAH3297.1; -.
DR EMBL; D00749; BAA00646.1; -.
DR EMBL; D00747; BAA00646.1; JOINED.
DR EMBL; D00748; BAA00646.1; JOINED.
DR PIR; A39016; A39016.
DR HSSP; P01607; IREI.
DR GENEW; HGNC:1695; CD7.
DR MIM; 186820; -.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0005866; C:plasma membrane; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0042110; P:T-cell activation; TAS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; NMS.
DR InterPro; IPRO0710; Ig-like.
DR InterPro; IPRO0359; IG.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW T-cell; Signal; Immune response; Antigen; Transmembrane; Glycoprotein; Immunoglobulin domain; Receptor; Lipoprotein; Palmitate; Repeat.
FT SIGNAL 1 25
FT CHAIN 26 240
FT DOMAIN 26 180 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 181 201 PROBABLE.
FT DOMAIN 202 240 CYTOPLASMIC (PROBABLE).
FT DOMAIN 26 130 IG-LIKE.
FT DOMAIN 145 180 4 x 9 AA TANDEM REPEATS, POTENTIAL SPACER FUNCTION.
FT DISULFID 35 142 POTENTIAL.
FT DISULFID 48 114 POTENTIAL.
FT LIPID 198 198 S-palmitoyl cysteine.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . . ).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . . ).
FT REPEAT 145 153 1.
FT REPEAT 154 162 2.
FT REPEAT 163 171 3.
FT REPEAT 172 180 4.
SO SEQUENCE 240 AA; 25409 MW; EBBCE08279552108 CRC64;

Query Match 11.7%; Score 316.5; DB 1; Length 240;
Best Local Similarity 41.7%; Pred. No. 1.6e-13;
Matches 90; Conservative 23; Mismatches 54; Indels 49; Gaps 11;

```

Db	14	LALARGLPGLAAGVEVQSSPHCTTVPVGA-----SVNICTSTGGRGRIYRLQLGPROQP	67
Qy			
Db	361	DIAYEWSNGCPENNYKTTTPVLDSDGSFFLYSKLTLYDKSRMOQGN--VSSCSYMHIALH	418
Qy			
Db	68	DI-LIYEDGVVPTDRFRGRRI-DPSGS---QDNLTITTHRLQLSDPTGYTCQAITV--	120
Qy			
Db	419	NHYQKSLSPGLQD--ETCAEAGCGELDGLWTTPPPASALPAPPTSSALPDPTQAS	476
Qy			
Db	121	NYSSGLVLVTEEOSQGMHRCSDA-----PPASALPAPPTSSALPDPTQAS	168
Qy			
Db	477	ALPDPAPASALPALAVISFLGLGIGVACVLARTR	512
Qy			
Db	169	ALPDPAPASALPALAVISFLGLGIGVACVLARTQ	204
Qy			
RESULT 47			
HVC2_HETFR			
ID	HVC2_HETFR	STANDARD;	PRT; 438 AA.
AC	P23085;		
DT	01-NOV-1991 (Rel. 20, Created)		
DT	01-NOV-1991 (Rel. 20, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DS	Ig heavy chain C region (Clone 12022) (Fragment).		
OS	Heterodontus francisci (Horn shark).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;		
OC	Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;		
NCBI	taxid=7792;		
NCBI	[1]		
RN	SEQUENCE FROM N.A.		
RP	MEDLINE=88328985; PubMed=138109;		
RX	Kokubu F., Hinde K., Litman R., Shambloet M.J., Litman G.W.;		
RA	"Complete structure and organization of immunoglobulin heavy chain		
RT	constant region genes in a phylogenetically primitive vertebrate.";		
RL	EMBO J. 7:1979-1988(1988).		
CC	-I-SIMILARITY: Contains 4 immunoglobulin-like domains.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way that		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
CC	EMBL; X07784; CAA30617.1; -.		
DR	PIR; S00980; HVRK2.		
DR	HSSP; P01842; 7PAB.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003597; Ig_c1.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	Pfam; PF00047; Ig; 4.		
DR	SMART; SMO0407; IgC1; 2.		
DR	PROSITE; PSS0835; IG_LIKE; 4.		
DR	PROSITE; PSS00290; IG_MHC; 3.		
KW	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.		
FT	NON_TER	1	
FT	DOMAIN	5	101
FT	DOMAIN	108	202
FT	DOMAIN	213	305
FT	DOMAIN	315	415
FT	CARBOHYD	166	166
FT	CARBOHYD	200	200
FT	CARBOHYD	245	245
FT	CARBOHYD	275	275
FT	CARBOHYD	374	374
FT	CARBOHYD	411	411
FT	CARBOHYD	415	415
FT	CARBOHYD	425	425
FT	SEQUENCE	438 AA;	47904 MW;
Query Match	11.5%;	Score 312;	DB 1;
Best Local Similarity	25.4%;	Pred. No. 6;	Length 438;

Matches 113; Conservative 67; Mismatches 155; Indels 110; Gaps 17;

QY 23 TQGNKV---VLGKGGDTVELTCAQKKSIGFHWKNSNOIKLNGSGFLTGPSKLNDR 79  
 DB 49 TGLGKTYPSVLNKKGT-----YTRSSQULTTRES-----VGSSKTYCE 86  
 QY 80 ADSRSLMDQGNFPLIKNLIKIEDSDTYICEVEDQKEEVLVFGLTANSDDL----- 133  
 DB 87 VARGESLW-----IKELIDCKGDIYPPVILITLSSSEETIRRA 126  
 QY 134 -----LOGGSLT-----LESPPGSSPSVOCRSRPGNKIOGKTLVSQLELD 178  
 DB 127 TVLCSIIDPHPSITSMKDGQPMDSGFVTSPTCEVNG---NFSATSLTLVPAGEWPS 182  
 QY 179 SGMTCTVLQNGOKVFXI--DIVPCPAPBPKSCDTHHCPELLGSPSLPPKXKDL 236  
 DB 183 NTVYTQVHAQETQSRNITGSOVPS-----IGDPVTKLPPSLIEQVL 226  
 QY 237 MISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTRPREQY-----NSTYRVSVL 291  
 DB 227 -LEATVTLTCV---VSNAPYGVNVSM-----TQKKPLKSEIADVGEBSDSVISTV 274  
 QY 292 TVLHODMLNGKEYKCVSNKALPAPIEKTISKAKGP-REPOY-VTLPPSRDELTKNOYS 349  
 DB 275 NISTQAMLSGAEFYCVVSHQDLPTPLRASIHKEEVKDLREPFVSVLLPPAEDVSAQRFLS 334  
 QY 350 LTVCLVGFPSDIAVEMESNGOPEN--NYKTPPVLDSCG-SPEFLYSKLTVDKSRQGN 406  
 DB 335 LTVCLVGFSPRELFIMVTWNDKSVNGVKNTEVMAENDNSFFIYSLLSIAAEWASGA 394  
 QY 407 VFSCSVMEHALNHNHYTKSLSPG 431  
 DB 395 SYSCVGVCHRAIPKIKINRTVNKSSG 419

RESULT 48  
 MUC\_CHICK  
 ID MUC\_CHICK STANDARD; PRT; 446 AA.  
 AC P01675;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig mu chain C region.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE OF 80-446 FROM N.A.  
 RX MEDLINE=63299221; PubMed=6310496;  
 RA Dahan A., Reynaud C.-A., Weill J.-C.;  
 RT "Nucleotide sequence of the constant region of a chicken mu heavy  
 chain immunoglobulin mRNA."  
 RL Nucleic Acids Res. 11:5381-5389(1983).  
 RN [2]  
 RP SEQUENCE OF 1-79 FROM N.A.  
 RA Weill J.-C.;  
 RL Submitted (MAR-1986) to the EMBL/Genbank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).

CC EMBL: X01613; CAA25762.1; --  
 DR PIR: A02170; MHCH.  
 DR HSSP: P01857; IFCL.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003597; Ig\_c1.

DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig\_3.  
 DR SMART: SM00407; IGc1. 1.  
 DR PROSITE: PS00835; IG\_LIKE. 4.  
 DR PROSITE: PS00290; IG\_MHC; 3.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON TER 1 105  
 FT DOMAIN 1 105 CH1.  
 FT DOMAIN 106 209 CH2.  
 FT DOMAIN 210 316 CH3.  
 FT DOMAIN 317 427 CH4.  
 FT DOMAIN 428 446 C-TERMINAL REGION.  
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN)  
 FT DISULFID 27 85 (BY SIMILARITY).  
 FT DISULFID 135 190 BY SIMILARITY.  
 FT DISULFID 237 296 BY SIMILARITY.  
 FT DISULFID 284 284 INTERCHAIN (WITH A HEAVY CHAIN)  
 FT DISULFID 344 406 (BY SIMILARITY).  
 FT DISULFID 445 445 INTERCHAIN (WITH A HEAVY CHAIN)  
 FT CARBOHYD 45 45 (BY SIMILARITY).  
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 158 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 446 AA; 48173 MW; 3CB0CE108949BD17 CRC64;

Query Match 11.5%; Score 310.5; DB 1; Length 446;  
 Best Local Similarity 26.2%; Pred. No. 8.7e-13;  
 Matches 119; Conservative 67; Mismatches 179; Indels 89; Gaps 19;

QY 28 VLGKGGDTV-ELTCTA--SOKKSIQFHWKNSNOIKLNGSGFLTGPSKLNDRADSR 84  
 DB 13 LVLCSPSBSVTVGCAADPQSPSIAFTWPFDSNNSVSGM--DIVPKYISGPPYAVSR- 69  
 QY 85 SLMDQGNFPLIKNLIKIEDSDTYICEVEDQKEEVLVFG---LTANSDDLQGSULT 140  
 DB 70 -----IQNQGEGKEKQPFRCRAHPRGNVSVNPGPIPTPN-----GIPLF 113  
 QY 141 LTLSP-----PSSSSVQC-----RSPR-----GKXIQQGK 167  
 DB 114 VTMHPSPSEDEFGPPRNASILCQTRGRRPTEVTYKNGSPVAATAATTATVGPVVAES 173  
 QY 168 TLVSQLELDQSGTCTVLQNGKVEFKIDIVPCPAPBPKSCDTHHC---PELLGSPS 224  
 DB 174 RISVTESEMDGATFSCV-----EGEMRNTSKRMCGLEPVLVQODIA 216  
 QY 225 VLPFPKPKDITLMSRTEPVTGVVDVSHEDP-EVKFMYVDGVEVHNAKTRPREQVNS 283  
 DB 217 IRVITPSPVD-IFISKATLTCRVSNMVMNADGLEVSM-WKEKGGLLEVALGK-RVLQNG 273  
 QY 284 TYRVSVTLVTHQDMLNGKEYKCVSNKALPAPIEKTISKAK-GOPREPOVYTLPPSRDE 342  
 DB 274 LYTVDGVAIVCASEMDGDGDGVCKVNHDPDLFPMEKKRKTKASVAPRPSVVPFPPTQ 333  
 QY 343 LTKNO-VSLTCLVKGFYPSDIAVEMESNGO--PENNYTTTPVLDG--DGSPEFLYSKLT 397  
 DB 334 LKNGRLSVTCAQGFNPPLFLVRMRNGEPLPQOSVTSAPMAENPENESTVAIVSLGV 393  
 QY 398 DKSRMQGNVPSCVMEHALNHNHYTKSLSPG 431  
 DB 394 GAEEWGAGNVYTCVLGHALPLQLAQKSVDBASG 427

RESULT 49  
 HVCN\_HETFR  
 ID HVCN\_HETFR STANDARD; PRT; 461 AA.  
 AC P23088;  
 DT 01-NOV-1991 (Rel. 20, Created)



```
DR EMBL: J00221; AAB59396.1; ALT_INIT.
DR PIR: B2360; B2360.
DR HSSP: P01810; 2FBJ.
DR GeneW: HGNC:5479; IGHA2.
DR MIM: 147000; -.
DR GO: GO:0005624; C:membrane fraction; NAS.
DR GO: GO:0003823; F:antigen binding; TAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig-cl.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS50835; IG_LIKE; 3.
DR PROSITE: PS00290; IG_MHC; 2.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
FT NON_TER 1 98
FT DOMAIN 112 207
FT DOMAIN 115 317
FT DISULFID 26 85
FT DISULFID 101 101
FT DISULFID 109 109
FT DISULFID 110 167
FT DISULFID 134 191
FT DISULFID 169 179
FT DISULFID 179 179
FT DISULFID 237 300
FT DISULFID 339 339
FT CARBOHYD 47 47
FT CARBOHYD 92 92
FT CARBOHYD 131 131
FT CARBOHYD 205 205
FT CARBOHYD 327 327
FT VARIANT 93 93
FT VARIANT 102 102
FT VARIANT 279 279
FT VARIANT 296 296
FT VARIANT 326 326
FT VARIANT 335 335
SQ SEQUENCE 340 AA; 36508 MW; 98922700756F3276 CRC64;

Query Match 10.9%; Score 294.5; DB 1; Length 340;
Best Local Similarity 29.2%; Pred. No. 6.5e-12;
Matches 103; Conservative 50; Mismatches 143; Indels 57; Gaps 18;

QY 122 VFGLTANSDTH-----LLOG-----QSLLTLESPPGSSPVOCRS-PRGNIGGKTK 168
DB 8 VFPLSLDSTFGQDGNVYVACLVGQFPQEPRLSVTWS-----SGQVNTARNPPSSDAGDLY 64
QY 169 LSVSGL-----QDSGTWCTVLONQKVEFKIDI-VPCPAPEPKSCDKHTCPBELGG 222
DB 65 TTSSQTLPRATQCPDGKSVTCHVKHYTNPSQ---DVTVPFVPPPPCC-----C-----H 111
QY 223 PSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNYYVDGVEVNAKTPREEQVN 282
DB 112 PRSLTRPALED--LLLGSEANLICTTLGL--RDASGATFTWTPSSGK--SAVQGPPEPRLDC 167
QY 283 STYRVVSVLTFLVLTQHDMNGEKYCKVSNKALPAPIEKTSKAGQPREPOVYTLPPSRDE 342
DB 168 GCSVSVSVLRGCAGQPNNHGETFTCTAAHPELKTPLRANITKS--GNFRPREVHLPPPSSE 226
QY 343 LTKNQ-VSLTCLVKGFPYSDIAVEMESNQ--PENNYKTPPYLD--SDG--SFFLYSKLT 396
DB 227 LALNELVTLTCLARGSPKDVLRVMLQSGQELPREKYLTVASRQEPSOGTTTFAVTSILR 286
```

```
QY 397 VDKSRQGNVSCSYMEALHNHYTKSLSPG-----LQIDETC 438
DB 287 VAAEDWKKDDTSCVNGHEALPLATFQKTIIDLAKGPTHVNVSVMAEVDGTC 339

RESULT 51
HVC5_HETFR
ID HVC5_HETFR STANDARD; PRT; 438 AA.
AC P23087;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain C region, secreted form (clone 3050).
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Heterodontidae; Heterodontiformes;
OC Heterodontidae; Heterodontus.
OX NCBI_Taxid=7792;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=85328985; PubMed=3138109;
RA Kokubu F., Hinds K., Litman R., Shambloot M.J., Litman G.W.;
RT "Complete structure and organization of immunoglobulin heavy chain
constant region genes in a phylogenetically primitive vertebrate.";
RL EMBO J. 7:1979-1988(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X07781; CAA30613.1; ALT_INIT.
CC PIR: S01853; HVRKCS.
CC HSSP: P01842; 7FAB.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig-cl.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_4.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS50835; IG_LIKE; 4.
DR PROSITE: PS00290; IG_MHC; 3.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1 99
FT DOMAIN 1 100
FT DOMAIN 100 205
FT DOMAIN 206 308
FT DOMAIN 309 418
FT CARBOHYD 164 164
FT CARBOHYD 200 200
FT CARBOHYD 245 245
FT CARBOHYD 275 275
FT CARBOHYD 374 374
FT CARBOHYD 411 411
FT CARBOHYD 415 415
FT CARBOHYD 425 425
SQ SEQUENCE 438 AA; 47920 MW; 9BDF37C3BBD7D15B CRC64;

Query Match 10.9%; Score 294; DB 1; Length 438;
Best Local Similarity 27.1%; Pred. No. 9.6e-12;
Matches 115; Conservative 65; Mismatches 153; Indels 92; Gaps 21;

QY 53 WKNNSNIRKILNGQSFLLTGPSKLDNRADSRSLWDQNGFPLIINKLKIEDSDTYICEV- 111
DB 41 WKDNEPIITG-----LKTYPVLLNKKGYTDS-----SOLLTLESSESSKIY-CEVR 88
QY 112 -----EDQKEEVLIVFG--LTRANSDFHLLQGSLLTL-- 142
DB 89 RGESWIKEIPPOCKDKXHPVILVLTQSSSEELTSRRFAVLCSIIDFH---PESTIVSWL 145
```

[illegible]

ID	ALC_RABIT	STANDARD;	PRT;	299 AA.
AC	P01879;			
DT	21-JUL-1986	(Rel. 01, Created)		
DT	21-JUL-1986	(Rel. 01, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	IG alpha chain C region (Fragment).			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_taxid=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84144059; PubMed=6322114;			
RA	Knight K.L., Marens C.L., Stockloa C.M., Schneiderman R.D.;			
RT	"Genes encoding alpha-heavy chains of rabbit IgA: characterization of			
RT	cDNA encoding IgA-g subclasses alpha-chains.";			
RL	Nucleic Acids Res. 12:1657-1670(1984).			
CC	-1- FUNCTION: Ig alpha is the major immunoglobulin class in body			
CC	secretions. It may serve both to defend against local infection			
CC	and to prevent access of foreign antigens to the general			
CC	immunologic system.			
CC	-1- MISCELLANEOUS: This immunoglobulin belongs to the IgA-G subclass.			
CC	It was isolated from a rabbit homozygous FOR A2, N80, DE12, 15,			
CC	F71, G75 heavy chain haplotype.			
CC	-1- SIMILARITY: Contains 2 immunoglobulin-like domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .			
CC	-----			
DR	EMBL; X00353; CAA25100.1; -.			
DR	PIR; A02174; AHRB.			
DR	HSP; P01857; 1FCL.			
DR	InterPro; IPR007110; IG_1like.			
DR	InterPro; IPR003597; IG_c1.			
DR	InterPro; IPR003006; IG_MHC.			
DR	Pfam; PF00047; Ig; 2.			
DR	SMART; SM00407; IGL1; 2.			
DR	PROSITE; PS50835; IG_LIKE; 2.			
DR	PROSITE; PS00290; IG_MHC; 2.			
FW	Immunoglobulin domain; Immunoglobulin C region; Repeat.			

FT	NON TER	1	1	IG-LIKE 1.
FT	DOMAIN	71	167	IG-LIKE 2.
FT	DOMAIN	174	276	
SO	SEQUENCE	299 AA;	32256 MM;	2512FE3FE2E9A223 CRC64;

Query Match	10.8%;	Score 293;	DB 1;	Length 299;
Best Local Similarity	31.7%;	Pred. No. 6.9e-12;		
Matches	78;	Conservative	42;	Mismatches 106;
			Indels	20;
			Gaps	10;

```

Qy      200  VPCAPRPKSCDKHNTCP-----ELIG--PSVLPFRPKMDTMISTREVTGVVUDVS 252
Db      41  LPVFPFCCPCCANSCCTSPSSSRLLIGCCPSSLYQRPDLG-ILIGDABLTTLSGLK 99
Qy      253  HEDEVKFNMYVDGVEVHNAKTKREBOYNSTYRVSVLTGLVHDMWLNKEXKCKVSNKA 312
Db      100  NPEBAV-FTW--EPTNGNBPVQGRADRLSGCYSVSSLTSSAETWAKRTFTCTVTHPE 156
Qy      313  LPA-PIEKTSKAKGPREPOUYTLPRSRDELTKN-QVSLTCLVKGFRPSDIAYEWESNG 370
Db      157  IDSSGLTATTSIRGVNTP--PQVHLPRPSEELALNEBOVTLTCLVRGFSRKVLVSMRHQG 214
Qy      371  Q--PENN---KTPRVLVDSGFFLYSKULVDKSRQOQGVFCSSVHNEALHNHYTQKS 425
Db      215  QEVEDSEFLVWKSMPESSQDKATVATSLLRVRPMDNQGQDTYCSVGNHEGLAHFNQKT 274
Qy      426  LSLSPG 431
Db      275  IDRLAG 280

```

RESULT 53	ID	SV	STANDARD;	PRT;	370 AA.
AC	P23084;				
DT	01-NOV-1991 (Rel. 20, Created)				
DT	01-NOV-1991 (Rel. 20, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Ig heavy chain C region (Clone 6125) (Fragment).				
OS	Heterodontus francisci (Horn shark).				
OS	Eulachya; Metacoa; Chordata; Craniota; Vertebrata; Chondrichthyes;				
OC	Elasmobranchii; Galeomorphii; Heterodontoidae; Heterodontiformes;				
OC	Heterodontidae; Heterodontus.				
ON	NCBI_Taxid=7792;				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=68328985; PubMed=3138109;				
RA	Kokubu F., Hinds K., Litman R., Shambloct M.J., Litman G.W.;				
RT	"Complete structure and organization of immunoglobulin heavy chain				
RT	constant region genes in a phylogenetically primitive vertebrate.",				
RL	EMBO J. 7:1979-1988(1988).				
CC	- I SIMILARITY: Contains 3 immunoglobulin-like domains.				
CC					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
DR	EMBL; X07783; CAA30616.1; -				
DR	PIR; S01851; HVKKS.				
DR	HSSP; P01857; 1FCL.				
DR	InterPro; IPR007110; Ig_1like.				
DR	InterPro; IPR003597; Ig_c1.				
DR	InterPro; IPR003006; Ig_MHC.				
DR	Pfam; PF00047; Ig_3.				
DR	SMART; SM00407; Igcl; 2.				
DR	PROSITE; PSS0835; IG_LIKE; 3.				
DR	PROSITE; PSS00290; IG_MHC; 3.				
DR	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.				
TM	NON TER				
TM	DOMAIN				
TM	40 134				
TM	1 1				
TM	IG-LIKE 1.				

```

CC FT DOMAIN 145 237 IG-LIKE 2.
FT DOMAIN 247 347 IG-LIKE 3.
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 370 AA; 40586 MW; 32833ABEFADDD0 CRC64;

Query Match 10.6%; Score 286.5; DB 1; Length 370;
Best Local Similarity 26.2%; Pred. No. 2.3e-11;
Matches 101; Conservative 59; Mismatches 155; Indels 71; Gaps 16;

QY 72 GSGKNDNRADSRSLM-----DQGNFPLIKNLKIEDSD-----TYICEVEQKE 117
DB 11 GSKKICEVRGESLWIKELPDCKGIVHPVTYILTGTSSEETSSRFATVCSIIDFHE 70
QY 118 VOLVLEGLTAN--SDPHLLQGGSLTLTLESPPGSSPVOCSPRGKNIQGGKTLVSQLE 175
DB 71 -----AIVNMKIDGHPMESGFT-----SACETNG-----NFSATSRITVPARE 111
QY 176 LODSGTWCTVTLQNGKKEFKIDIVCPAPEPKSCDKHTCPBELLGSPVFLPPPKDPT 235
DB 112 WFTNTVYTCQVSHQATQGRNI-----TGSPPDSECNH-----PAIKLLPPIEGV 157
QY 236 LMSRPEPTCVVDVSHEDPEYKFNWYVDGVEVNAKTKPREQY-----NSTYVSV 290
DB 158 L-LEAVTLTLCV---VSNAFYGVNSW-----TQEQKPLKSEIAVOPGEDPDSVIST 205
QY 291 LTVLHODMNGKEYKCKVSNKALPAPIETKISK-ANGQPREPOVYTLPPSRDELTKNQ-V 348
DB 206 VDISQAWLSEAVFYCVVSHODLPTLRDSIHKEAKDLREPEVSULPRAEISMERL 265
QY 349 SLTCLVKGFFPSDIAVESNGQPEN--NYKTPPVLDSD-GSFPLYSKLTVDKSRWQ 405
DB 266 SLTCLVGRGSPREIFVKMTYNDKSNVPGVYKNTVMAENDKSFYISLSTIAEEMWAG 325
QY 406 NVPSCVNHLEALHNHTOKSLSPG 431
DB 326 ASYSCVGHGHAIPLKIINRTVKNSSG 351

RESULT 54
HVC3_HETFR STANDARD; PRT; 393 AA.
ID HVC3_HETFR STANDARD; PRT; 393 AA.
AC P23086;
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
OS Ig heavy chain C region (Clone 6121) (Fragment).
OC Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;
OC Heterodontidae; Heterodontus.
OC NCBI_Taxid=7792;
OK RN
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=88328985; PubMed=3138109;
RA Kokubu F., Hinde K., Litman R., Shambloot M.J., Litman G.W.;
RT "Complete structure and organization of immunoglobulin heavy chain
RL EMBL J. 7:1979-1988 (1988).
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL: X07782; CAA30615.1; -.
DR PIR: S01852; HVRK1.
DR HSSP: P01842; 7FMB.
DR InterPro: IPR007110; Ig-1-like.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_3.
DR SMART, SM00407; IgC1_2.
DR PROSITE: PS00835; Ig_LIKE, 3.
DR PROSITE: PS00290; Ig_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
FT NON TER 1 1
FT DOMAIN 63 157 IG-LIKE 1.
FT DOMAIN 158 250 IG-LIKE 2.
FT DOMAIN 270 370 IG-LIKE 3.
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 393 AA; 43081 MW; 4E44D076972F18B5 CRC64;

Query Match 10.4%; Score 281; DB 1; Length 393;
Best Local Similarity 26.4%; Pred. No. 5.7e-11;
Matches 115; Conservative 69; Mismatches 161; Indels 90; Gaps 21;

QY 23 TQGNKV---VLGKGG---DTVELTCTASOKKSIQPHMKNSNQIKLNGSGFLTKGPSKL 76
DB 4 TTGLKTPSVLKKKGTYYOSSQLTTTSEF-----VQSSKI 38
QY 77 NDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEQKEVQLVFG--LTNSDTHLL 134
DB 39 YCEVVRGESVWIKI-----IPCKGDKVHPVTYILQSSSEETSSRFATVCSIIDFH-- 91
QY 135 QGQSLLT-----LESPPGSSPVOCSPRGKN--IQGGKTLVSQLELDSGTWCT 185
DB 92 -PESITVSWLKDQGMESGFVTSPT--C---GVNGTSPATSRALTVPAREMPTNKRYTQ 144
QY 186 VLQNGKKEFKIDI---VPCPAPKPSCKDKHTCPBELLGSPVFLPPPKDTLMSRT 241
DB 145 V--SHQGTQSKNRTIGSQVPC-----SCN-----DVIKLLPPISTQVL-LENT 185
QY 242 PEVTCVVDVSHEDPEYKFNWYVDGVEVNAKTKPREQYNSYRVSVTLVHODMNG 301
DB 186 VTLNCI---VSNAFYGVNSW-----TQEQKSLKSEIAVOPGEDADSVISTVNI STQAWMSG 239
QY 302 KEYKCKVSNKALPAPIETKISAKQOP-REPOV-YTLPPSRDELTKNOVSLTCLVKGFP 359
DB 240 ASFYCVVNHQDLPPTLRASIHKEEYVDLREPEVSILLSPAEDVSAQRFLSLTCLVKGFS 299
QY 360 SDIAVESNGQPEN--NYKTPPVLDSD-GSFPLYSKLTVDKSRWQGNVSCVYMHBA 416
DB 300 REIFVKMTINDKSNVPGVYKNTVMAENDNSYFISLSTIAEEMWAGASYCVVGHBA 359
QY 417 LHNHYTOKSLSPG 431
DB 360 IPLKIINRTVKNSSG 374

RESULT 55
ID ALCL GORGO STANDARD; PRT; 353 AA.
AC P20758;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig alpha-1 chain C region.
GN IGHA1.
OS Gorilla gorilla gorilla (Lowland gorilla).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Gorilla.
OX NCBI_TaxID=9595;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=89386006; PubMed=2506527;
RA Kawamura S., Omoto K., Ueda S.;
RT "Nucleotide sequence of the gorilla immunoglobulin alpha 1 gene.";
CC Nucleic Acids Res. 17:6732-6732(1989).
CC -1- FUNCTION: Ig alpha is the major immunoglobulin class in body
CC secretions. It may serve both to defend against local infection
CC and to prevent access of foreign antigens to the general
CC immunologic system.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15045; CAA33147.1; ALT_INIT.
CC HSSP; P01810; 2F8J.
CC InterPro; IPR007110; Ig-1-like.
CC InterPro; IPR003597; Ig_C1.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig_3.
CC SMART; SM00407; IgC1_2.
CC PROSITE; PS00835; IG_LIKE; 3.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 98
FT FT 125 220
FT DOMAIN 228 330
FT FT 14 14
FT DISULFID 26 85
FT FT 77 101
FT DISULFID 122 122
FT FT 123 180
FT FT 147 204
FT DISULFID 182 182
FT DISULFID 192 192
FT FT 250 313
FT DISULFID 352 352
FT FT 144 144
FT CARBOHYD 340 340
FT FT 353 AA; 37755 MW; 4820EBD02AC7514 CRC64;
SQ SEQUENCE
Query Match 10.3%; Score 279; DB 1; Length 353;
Best Local Similarity 28.0%; Pred No. 6.6e-11;
Matches 96; Conservative 41; Mismatches 126; Indels 80; Gaps 16;

```

```

DB 191 GCAEPNMHNGKFTCTAAVPSKPTLATLSKS -GNMFRPEVHLPPSEELALNELVLT 249
OY 352 CLVKGFPSPDIAVWESNGC--PENNYKTPPVLD-SDG--SEFLYSGLTYDKRMOGN 406
DB 250 CLARGFSPKDVLMWLGOSQELPREKYLTVASRQSPQSTTFVATSLRVAADWKGD 309
OY 407 VFSCVWHEALHNHYTOKSLSPG-----IQLDQTC 438
DB 310 TFSQVGHLEALPLAFQKIDRLAGKPTHVNVSVMAEVDTC 352
RESULT 56
ALCI_HUMAN STANDARD; PRT; 353 AA.
AC P01876;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ig alpha-1 chain C region.
CN IGHAI1
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84130179; PubMed=6421489;
RA Flanagan J.G., Lefranc M.-P., Rabbitts T.H.;
RT "Mechanisms of divergence and convergence of the human immunoglobulin
RT alpha 1 and alpha 2 constant region gene sequences.";
RL Cell 36:681-688(1984).
RN [2]
RP SEQUENCE (MYELOMA PROTEIN BUR), AND DISULFIDE BONDS.
RX MEDLINE=79151016; PubMed=107164;
RA Putnam F.W., Liu Y.-S.V., Low T.-L.K.;
RT "Primary structure of a human IgM1 immunoglobulin. IV. Streptococcal
RT IgM1 protease, digestion, Fab and Fc fragments, and the complete
RT amino acid sequence of the alpha 1 heavy chain.";
RL J. Biol. Chem. 254:2865-2874(1979).
RN [3]
RP SEQUENCE (MYELOMA PROTEIN TRO).
RX MEDLINE=76023781; PubMed=809311;
RA Kratzin H., Altevogt P., Ruban E., Korot A., Staroscik K.;
RA Hilschmann N.;
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.),
RT II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT structure of the complete IgA-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
RN [4]
RP DISULFIDE BONDS.
RX MEDLINE=80114124; PubMed=393607;
RA Yang C.-Y., Kratzin H., Goltz H., Hilschmann N.;
RT "Rule of antibody structure. Primary structure of a human monoclonal
RT IgA-immunoglobulin (myeloma protein Tro). VII. Purification and
RT characterization of the disulfide bridges.";
RL Hoppe-Seyler's Z. Physiol. Chem. 360:1919-1940(1979).
RN [5]
RP REVIEW.
RX MEDLINE=91054387; PubMed=2241915;
RA Kerr M.A.;
RT "The structure and function of human IgA.";
RL Biochem. J. 271:285-296(1990).
CC -1- FUNCTION: Ig alpha is the major immunoglobulin class in body
CC secretions. It may serve both to defend against local infection
CC and to prevent access of foreign antigens to the general
CC immunologic system.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```



```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00220; AAC82528.1; ALT_INIT.
DR PIR; A23360; ALHU.
DR PDB; 1IGA; 15-JUN-99.
DR Genew; HGNC:5478; IGH1.
DR MIM; 146900; -.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KM Polymorphism; 3D-structure; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 98
FT 125 220 IG-LIKE 1.
FT 128 330 IG-LIKE 2.
FT DISULFID 14 14 IG-LIKE 3.
FT 26 85 INTERCHAIN (WITH LIGHT CHAIN).
FT 77 101 INTERCHAIN (WITH HEAVY CHAIN).
FT 122 122 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 123 180 OR 123-182 (IN REF. 4).
FT DISULFID 147 204 INTERCHAIN (WITH HEAVY CHAIN) (OR 180, IN
FT DISULFID 182 182 REF. 4).
FT DISULFID 192 192 INTERCHAIN (WITH HEAVY CHAIN OF ANOTHER
FT DISULFID 192 192 SUBUNIT) (PROBABLE).
FT DISULFID 250 313 INTERCHAIN (WITH J CHAIN).
FT 352 352 O-LINKED.
FT CARBOHYD 105 105 O-LINKED.
FT CARBOHYD 111 111 O-LINKED.
FT CARBOHYD 113 113 O-LINKED.
FT CARBOHYD 119 119 O-LINKED.
FT CARBOHYD 121 121 O-LINKED.
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .).
FT VARIANT 176 176 E -> D (in dBSNP:1407).
FT 163 165 /FTId=VAR_014602.
FT 176 176 TFS -> PST (IN REF. 2).
FT 190 190 E -> S (IN REF. 3).
FT 227 227 P -> B (IN REF. 3).
FT 231 231 R -> H (IN REF. 3).
FT 231 231 H -> R (IN REF. 3).
FT 290 290 T -> E (IN REF. 3).
SQ SEQUENCE 353 AA; 37654 MW; EBA11ECB7E85DB21 CRC64;
Query Match 10.3%; Score 277.5; DB 1; Length 353;
Best Local Similarity 27.2%; Pred. 8.3e-11;
Matches 104; Conservative 51; Mismatches 153; Indels 75; Gaps 18;

```

```

QY 313 LAPLEKTSKAGQPREQVYTLPSRDELTKNO-VSLTCLVKGFEYSDIAVEMESNGQ 371
DB 211 SKTPPLATLSKS-GVTFREYVHLPPSEELNLNLTCTLARGSPDVVVRMQSGQ 269
QY 372 --PENNYKTPPEVL-DSD--SFPLYSKLYDKSRMOQNVSCSYMHEALNNHYOKSL 426
DB 270 ELPREKYLTMASROPSQCTTFATVSLIRVAEDMKGDFTSCVNGHEALPLAFQKTI 329
QY 427 SLSPG-----LQDETC 438
DB 330 DRLAGKPTHVNVVMAAEVDGTC 352
RESULT 57
MUCM_ICTPU
ID MUCM_ICTPU STANDARD; PRT; 481 AA.
AC P23735;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ig mu chain C region membrane-bound form.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OK NCBI_TaxId=9998;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90384824; PubMed=2119496;
RA Wilson M.R., Marcuz A., van Ginkel F., Miller N.W., Clem L.W.,
RA Middleton D., Warr G.W.;
RT "The immunoglobulin M heavy chain constant region gene of the channel
RT catfish, Ictalurus punctatus: an unusual mRNA splice pattern produces
RT the membrane form of the molecule."
RL Nucleic Acids Res. 18:5227-5233(1990).
CC -I- MISCELLANEOUS: During differentiation, B lymphocytes switch from
CC expression of membrane-bound Igm to secretion of Igm. The mu
CC chains of membrane and secreted Igm differ in their C-terminal
CC segments.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL; X52617; -, NOT_ANNOTATED_CDS.
CC HSSP; P01857; 1FC1.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG_c1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 1.
DR PROSITE; PS50835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 2.
KM Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KM Transmembrane.
FT NON_TER 1 1
FT DOMAIN 1 101 CH1.
FT 102 203 CH2.
FT DOMAIN 204 304 CH3.
FT 305 434 CH4.
FT TRANSMEM 459 478 POTENTIAL.
FT 479 481 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 14 14 INTERCHAIN (WITH A LIGHT CHAIN)
FT DISULFID 14 14 (BY SIMILARITY).
FT DISULFID 26 85 BY SIMILARITY.
FT DISULFID 129 188 BY SIMILARITY.
FT DISULFID 233 284 BY SIMILARITY.
FT DISULFID 329 391 BY SIMILARITY.

```

```

FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 481 AA; 53337 MW; 668E784E181C71BA CRC64;

Query Match 10.2%; Score 276.5; DB 1; Length 481;
Best Local Similarity 23.1%; Pred No. 1.4e-10;
Matches 113; Conservative 72; Mismatches 156; Indels 149; Gaps 23;

QY 37 VELTCTASOKS---IQFHKNKNOIKILGNOSFLTKGPKSLNDRSRSLMDQGNP 93
DB 22 VTLGCVTRDLASADGLSFIWKDAS-----GSLT-----DIVGY 56
QY 94 LI-----IKNLKIEDSD-----TYICEV-----DQKEVQLLVFGLT 126
DB 57 AVQATGTYTSVSHVVKASDMWGNKKFTCEVKNGLGSKDASLQKEVERELHASLL---LT 113
QY 127 ANSDTHLLOGSLTTLTLESPGSSPVOCRSPRG-----KNL---QGG 166
DB 114 TPTQTEINDGRTPTVCLATP-----FSFKSHPTKMLEKTIDISNKKENIVSQNG 164
QY 167 KTLVSQLELODSGTWT-----CTVLONOKVAFKIDVPCAPAPKSCDKHTCPBL 219
DB 165 NFTAISVLELSAS-EMTSSTSPVKCEFOQKHNHFKEASVAPGDTKQO-----VK 214
QY 220 LGGPVSFLPPKPKOTLMISRTPEYTCVAVVSHDEPEVK-PNVVVDGVEVNAKTRRE 278
DB 215 ITGPT-----EDILIKRAGQEC---RAEDTGFKSIKWLIGRREISSLSNL--- 259
QY 279 EQYNSTYRVSVLT-VLHODMLNGEKYKSNKALPAPIEK-TISKAKGPREQVYTL 336
DB 260 ----SSKTTVSIQTHIGPEEMWNGTEFICEVHEHFTQGYEVATYKRENGNPEPKVYL 315
QY 337 PPSRDELTKNOVSLTCLYKGFPSDIAYEWESN-----QGPENNYKTPPVLDSDGS 388
DB 316 AP--PESSGESVTLTCYVDYFPRKVAVSWLVNDQVEVVGVEQN---TTAVIDRNL 369
QY 369 FFLYKTLVVDKSRMOGVNFGSVVHVALNHTYQKSLSLSGLOLDETCAADGELDG 448
DB 370 FSVSYQLIKTADWNGSVFCLVYHESIKDCVRHISRIA-----KDSKTPTLVN 420
QY 449 LWTDPPTAS 458
DB 421 LTLTNPQSCS 430

RESULT 58
ALC_MOUSE ID ALC_MOUSE STANDARD; PRT; 344 AA.
AC P01878;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ig alpha chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NM_001001;
RP SEQUENCE FROM N.A. (MYELOMA ABE48 AND J558).
RX MEDLINE=81261947; PubMed=6790349;
RA "Auffray C., Nagelotte R., Sikorav J.-L., Heidmann O., Rougeon F.;
RT "Mouse immunoglobulin A: nucleotide sequence of the structural gene
RL for the alpha heavy chain derived from cloned cDNAs."
RL Gene 13:365-374 (1981).
RN RP SEQUENCE OF 1-213 (MOPC 47A).
RX MEDLINE=80049769; PubMed=115869;
RA Robinson E.A., Appella E.;
RT "Amino acid sequence of a mouse myeloma immunoglobulin heavy chain

```

```

RT (MOPC 47 A) with a 100-residue deletion."
RL J. Biol. Chem. 254:11418-11430 (1979).
RN [3]
RP SEQUENCE OF 1-254 AND 291-344 (M511).
RX MEDLINE=81054880; PubMed=6776528;
RA Robinson E.A., Appella E.;
RT "Complete amino acid sequence of a mouse immunoglobulin alpha chain
RT (MOPC 511)."
RL Proc. Natl. Acad. Sci. U.S.A. 77:4909-4913 (1980).
CC -1- FUNCTION: Ig alpha is the major immunoglobulin class in body
CC secretions. It may serve both to defend against local infection
CC and to prevent access of foreign antigens to the general
CC immunologic system.
CC -1- MISCELLANEOUS: The final C-region domain is deleted from Ref.2
CC chain. It was isolated from a myeloma protein that contains 1
CC light and 1 heavy chain per molecule, linked by a disulfide bond.
CC In contrast, normal mouse IgA molecules contain 2 light and 2
CC heavy chains and lack a light-heavy chain disulfide bond.
CC -1- MISCELLANEOUS: M511 sequence was isolated from a myeloma protein that
CC binds phosphorylcholine.
CC -1- MISCELLANEOUS: M511 sequence was compared with that of mouse MOPC
CC 47A, and a genetic mechanism for the deletion of the CH3 domain of
CC the mutant chain is proposed.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D11468; BAA0206.1; ALT_INIT.
DR PIR: A91479; AHMS.
DR HSSP: P01810; ZFBJ.
DR InterPro: IPR007110; Ig_1like.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS50835; Ig_LIKE; 3.
DR PROSITE: PS00290; Ig_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
FT NON TER 1
FT DOMAIN 6
FT DOMAIN 116
FT DOMAIN 219
FT DISULFID 26
FT DISULFID 76
FT DISULFID 114
FT DISULFID 138
FT CARBOHYD 38
FT CARBOHYD 101
FT CARBOHYD 331
FT CARBOHYD 331
FT VARIANT 2
FT VARIANT 18
FT VARIANT 67
FT VARIANT 73
FT VARIANT 112
FT VARIANT 112
FT VARIANT 135
FT VARIANT 141
FT VARIANT 141
FT VARIANT 168
FT VARIANT 212
FT VARIANT 213
FT VARIANT 235
FT VARIANT 255
FT VARIANT 295
FT VARIANT 295
FT VARIANT 301
FT VARIANT 301
FT VARIANT 329
FT VARIANT 331
SQ SEQUENCE 344 AA; 36676 MW; 3694CF9B919A98 CRC64;

Query Match 9.8%; Score 264.5; DB 1; Length 344;

```

	Best Local Similarity	26.7%: Pred. No. 5.4e-10:	
Marches	85;	Conservative	57; Mismatches 117; Indels 59; Gaps 15
Qy	136	GOSLTLTLESPPG-----	SSPSVQCRSPRGKIQGGKTLVSQLELDQSGTW 182
Db	45	GMDIT-TVNFPPALASGGRYTMNQUTLPAVEC--PEGSSVK-----	83
Qy	183	TCTVLQNOKKAEFKIDYVPCRPAPRPSCKRHTCPPELLGSPVFLFPYPRKDTLMSRTP 242	
Db	84	-CSVQSDSNPVQ-ELD-VNCSGPTPEPPPTIPSCQ-----PSLSLDRPALED-TLLGSDA 134	
Qy	243	EYTCVAVDVSHEDPEVKFMVYDVGVVNAKTRPREQYNST---YRVSVYLVTHQMDL 299	
Db	135	SITCTLNGIRNEGAV-FTW-----EPSTGKQAVQKQAVONSGCYSVSSVLPGCARKN 188	
Qy	300	NGKEYCKRVSNKALPAPRIEKTISKAKGQPREPOVYTLPPSRDELTKNQ-VSLTCLVKGFY 358	
Db	189	SGASRKCITYTHPE-SGLTGLGTIAKTVNFFPPQVHLPPPSBELANELLSTLCIVRATN 247	
Qy	359	PSDIVAVESNQ---PENNYKTTTPPVLDSDG--SFLLYSKLTVDKSRNQGVNFCGYM 413	
Db	248	PREVLVRLWMLHGNEELSPESYLVFEEPLKEGEGATVLYLTVSRVSAETWKQGDQYSCMVG 307	
Qy	414	HEALHNHTOKSLSPG 431	
Db	308	HEALPMNFTQKTIIDRLSG 325	
RESULT 59			
ID	SHS1_MOUSE	STANDARD;	PRT; 513 AA.
AC	P97797; Q08907; Q35924; Q08555; Q08556; P97796; Q08559; Q90X57; Q9WNTN4;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DE	Protein-tyrosine phosphatase non-receptor type substrate 1 precursor		
DE	(SHP substrate-1) (SHP-1) (Inhibitory receptor SHPS-1) (Signal-		
DE	antigen) (Brain Ig-like molecule with tyrosine-based activation		
DE	motifs) (Blt) (p64).		
GN	PENPS1 OR SHPS1 OR SHP OR MYD1 OR BIT.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RC	TISSUE=Brain;		
RC	MEDLINE=97223359; PubMed=9070220;		
RA	Yamao T., Matczak T., Amano K., Matsuda Y., Takahashi N., Ochi F.,		
RT	Fujioka Y., Kanuga M.;		
RT	"Mouse and human SHPS-1: molecular cloning of cDNAs and chromosomal		
RL	localization of genes.";		
RL	Biochem. Biophys. Res. Commun. 231:61-67(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANTS ALA-29; ARG-67; ARG-91;		
RP	THR-96; SER-128; PRO-194 AND ASN-224.		
RC	STRAIN=BALB/c; TISSUE=Brain;		
RC	MEDLINE=97230468; PubMed=9073552;		
RA	Ohnishi H., Kubota M., Sano S.-I.;		
RT	"Blt (Blt) maps to mouse chromosome 2.";		
RL	Genomics 40:504-506(1997).		
RN	[3]		
RP	SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), SEQUENCE OF 32-53 AND 422-433		
RP	(ISOFORM 2), VARIANTS ALA-29; ARG-91; THR-96 AND SER-128,		
RP	N-GLYCOSYLATION, AND TISSUE SPECIFICITY.		
RC	STRAIN=BALB/c; TISSUE=Brain, and Cerebellum;		
RC	MEDLINE=98012243; PubMed=9348339;		
RA	Comu S., Weng W., Olinsky S., Ishwed P., Mi Z., Hempel J., Mackins S.,		
RA	Lagenaur C.F., Narayanan V.;		
RT	"The murine p64 neural adhesion molecule is SHPS-1, a member of the		
RT	phosphatase-binding protein family.";		

RL J. Neurosci. 17:8702-8710 (1997).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), VARIANTS ALA-29; ARG-67;  
 RP PRO-74; ALA-83; 86-TYR-VAL-87; ILE-90; ARG-91; THR-96; ALA-114;  
 RP ILE-118; SER-128; PRO-194; ASN-224; PHE-351 AND ASP-365,  
 RP N-GLYCOSYLATION, PHOSPHORYLATION ON TYROSINE RESIDUES, INTERACTION  
 RP WITH PTPN6, AND TISSUE SPECIFICITY.  
 RP STRAIN=C57BL/6; TISSUE=Fetal thymus;  
 RC MEDLINE=98380500; PubMed=9712903;  
 RA Veilleux A., Thibodeau E., Lefcort S.,  
 RT "High expression of inhibitory receptor SHP-1 and its association  
 RL with protein tyrosine phosphatase SHP-1 in macrophages.",  
 RL J. Biol. Chem. 273:22719-22728 (1998).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3), VARIANTS ALA-29; ARG-67;  
 RP PRO-74; ALA-83; 86-TYR-VAL-87; ILE-90; ARG-91; THR-96; ALA-114;  
 RP ILE-118; SER-128; PRO-194; ASN-224; PHE-351 AND ASP-365, AND  
 RP N-GLYCOSYLATION.  
 RP STRAIN=129/SvJ and C57BL/6; TISSUE=Brain, and Liver;  
 RC MEDLINE=20053880; PubMed=10585853;  
 RA Sano S.-I., Ohnishi H., Kubota M.,  
 RT "Gene structure of mouse BIR/SHPS-1.",  
 RL Biochem. J. 344:667-675 (1999).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANTS ALA-29; ARG-67;  
 RP PRO-74; ALA-83; 86-TYR-VAL-87; ILE-90; ARG-91; THR-96; ALA-114;  
 RP ILE-118; SER-128; PRO-194; ASN-224; PHE-351 AND ASP-365.  
 RP STRAIN=11S, and 1S;  
 RC MEDLINE=21363810; PubMed=11471062;  
 RA Ehringer M.A., Thompson J., Conroy O., Xu Y., Yang F., Camniff J.,  
 RA Beeson M., Gordon L., Bennett B., Johnson T.E., Sikeja J.M.,  
 RT "High-throughput sequence identification of gene coding variants  
 RL within alcohol-related QTLs.",  
 RL Mamm. Genome 12:657-663 (2001).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT PRO-194.  
 RP TISSUE=Placenta;  
 RC Wang H., Chen Z., Ullrich A.,  
 RA "Epidermal growth factor-induced association of SHP2 with mouse SIRP-  
 RL alpha1.",  
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.  
 RN [8]  
 RP FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
 RC MEDLINE=90152134; PubMed=2303162;  
 RA Chuang W., Lagenaar C.F.,  
 RT "Central nervous system antigen P84 can serve as a substrate for  
 RL neurite outgrowth.",  
 RL Dev. Biol. 137:219-232 (1990).  
 RN [9]  
 RP N-GLYCOSYLATION, PHOSPHORYLATION BY JAK2 IN RESPONSE TO GROWTH  
 RP HORMONE, AND INTERACTIONS WITH JAK2 AND PTPN11.  
 RC MEDLINE=98175985; PubMed=9507023;  
 RX Stofega M.R., Wang H., Ullrich A., Carter-Su C.,  
 RA "Growth hormone regulation of SIRP and SHP-2 tyrosyl phosphorylation  
 RT and association.",  
 RL J. Biol. Chem. 273:7112-7117 (1998).  
 RN [10]  
 RP INTERACTION WITH CD47, AND TISSUE SPECIFICITY.  
 RC MEDLINE=99091586; PubMed=9872987;  
 RA Jiang P., Lagenaar C.F., Narayanan V.,  
 RT "Integrin-associated protein is a ligand for the P84 neural adhesion  
 RL molecule.",  
 RL J. Biol. Chem. 274:559-562 (1999).  
 RN [11]  
 RP INTERACTION WITH FGR.  
 RC MEDLINE=20130295; PubMed=10662797;  
 RX Gresham H.D., Dale B.M., Potter J.W., Chang P.W., Vines C.M.,  
 RA Lowell C.A., Lagenaar C.F., Willman C.L.,  
 RT "Negative regulation of phagocytosis in murine macrophages by the Src  
 RL kinase family member, Fgr.",  
 RL J. Exp. Med. 191:515-528 (2000).  
 CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Activates  
 as docking protein and induces translocation of PTPN6,

CC PTPN11 and other binding partners from the cytosol to the  
 CC plasma membrane. Supports attachment of cerebellar neurons, neurite  
 CC outgrowth and glial cell attachment. May play a key role in  
 CC intracellular signaling during synaptogenesis and in synaptic  
 CC function. Involved in the negative regulation of receptor tyrosine  
 CC kinase-coupled cellular responses induced by cell adhesion, growth  
 CC factors or insulin. Mediates negative regulation of phagocytosis,  
 CC mast cell activation and dendritic cell activation. CD47 binding  
 CC prevents maturation of immature dendritic cells and inhibits  
 CC cytokine production by mature dendritic cells (By similarity).  
 CC - SUBUNIT: Binds PTPN11 when tyrosine-phosphorylated, except in  
 CC macrophages, where it primarily binds PTPN6. Binds GRB2 vitro.  
 CC Binds FGR. Binds JAK2 irrespective of its phosphorylation status  
 CC and forms a stable complex. Binds SCAP1 and/or SCAP2. The  
 CC resulting complex recruits FVB. Binds PTK2B (By similarity).  
 CC - SUBCELLULAR LOCATION: Type I membrane protein.  
 CC - ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1; Synonyms=a;  
 CC IsoId=P97797-1; Sequence=Displayed;  
 CC Name=2; Synonyms=a; Large;  
 CC IsoId=P97797-2; Sequence=VSP\_007032;  
 CC Name=3; Synonyms=b, Small;  
 CC IsoId=P97797-3; Sequence=VSP\_007031;  
 CC - TISSUE SPECIFICITY: Highly expressed in cerebral cortex, brain,  
 CC spinal cord, cerebellum and spleen, and at much lower levels in  
 CC kidney, thymus, heart, lung and liver. Within the cerebellum,  
 CC highly expressed throughout the molecular layer, and in synaptic  
 CC glomeruli in the granule cell layer. Detected in neurons of the  
 CC hippocampus and dentate gyrus, and in olfactory bulb. Not detected  
 CC in Purkinje cells. Highly expressed in the plexiform layers, optic  
 CC fiber layer and the outer segments of the photoreceptor layer in  
 CC the retina. Highly expressed in macrophages. Isoform 3 is detected  
 CC at very low levels in all tissues tested.  
 CC - DEVELOPMENTAL STAGE: Highly expressed in the CNS of embryos from  
 CC day 7 to 17.  
 CC - PTM: N-glycosylated.  
 CC - PTM: Phosphorylated on tyrosine residues.  
 CC - SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.  
 CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL, D87967; BAA13520.1; -  
 CC DR EMBL, D87968; BAA13521.1; -  
 CC DR EMBL, D87969; BAA20376.1; -  
 CC DR EMBL, U89694; AAB92591.1; -  
 CC DR EMBL, AF072543; AAC24886.1; -  
 CC DR EMBL, AF072544; AAC24887.1; -  
 CC DR EMBL, AB024507; BAA89290.1; -  
 CC DR EMBL, AB024500; BAA89290.1; JOINED.  
 CC DR EMBL, AB024503; BAA89290.1; JOINED.  
 CC DR EMBL, AB024502; BAA89290.1; JOINED.  
 CC DR EMBL, AB024504; BAA89290.1; JOINED.  
 CC DR EMBL, AB024505; BAA89290.1; JOINED.  
 CC DR EMBL, AB024506; BAA89290.1; JOINED.  
 CC DR EMBL, AB024504; BAA89290.1; JOINED.  
 CC DR EMBL, AB024505; BAA89290.1; JOINED.  
 CC DR EMBL, AB024506; BAA89290.1; JOINED.  
 CC DR EMBL, AB024507; BAA89290.1; JOINED.  
 CC DR EMBL, AB024508; BAA89290.1; JOINED.  
 CC DR EMBL, AB024509; BAA89290.1; JOINED.  
 CC DR EMBL, AB024510; BAA89290.1; JOINED.  
 CC DR EMBL, AB024511; BAA89290.1; JOINED.  
 CC DR EMBL, AB024512; BAA89290.1; JOINED.  
 CC DR EMBL, AB024513; BAA89290.1; JOINED.  
 CC DR EMBL, AB024514; BAA89290.1; JOINED.  
 CC DR EMBL, AB024515; BAA89290.1; JOINED.  
 CC DR EMBL, AB024516; BAA89290.1; JOINED.  
 CC DR EMBL, AB024517; BAA89290.1; JOINED.  
 CC DR EMBL, AB024518; BAA89290.1; JOINED.  
 CC DR EMBL, AB024519; BAA89290.1; JOINED.  
 CC DR EMBL, AB024520; BAA89290.1; JOINED.  
 CC DR EMBL, AB024521; BAA89290.1; JOINED.  
 CC DR EMBL, AB024522; BAA89290.1; JOINED.  
 CC DR EMBL, AB024523; BAA89290.1; JOINED.  
 CC DR EMBL, AB024524; BAA89290.1; JOINED.  
 CC DR EMBL, AB024525; BAA89290.1; JOINED.  
 CC DR EMBL, AB024526; BAA89290.1; JOINED.  
 CC DR EMBL, AB024527; BAA89290.1; JOINED.  
 CC DR EMBL, AB024528; BAA89290.1; JOINED.  
 CC DR EMBL, AB024529; BAA89290.1; JOINED.  
 CC DR EMBL, AB024530; BAA89290.1; JOINED.  
 CC DR EMBL, AB024531; BAA89290.1; JOINED.  
 CC DR EMBL, AB024532; BAA89290.1; JOINED.  
 CC DR EMBL, AB024533; BAA89290.1; JOINED.  
 CC DR EMBL, AB024534; BAA89290.1; JOINED.  
 CC DR EMBL, AB024535; BAA89290.1; JOINED.  
 CC DR EMBL, AB024536; BAA89290.1; JOINED.  
 CC DR EMBL, AB024537; BAA89290.1; JOINED.  
 CC DR EMBL, AB024538; BAA89290.1; JOINED.  
 CC DR EMBL, AB024539; BAA89290.1; JOINED.  
 CC DR EMBL, AB024540; BAA89290.1; JOINED.  
 CC DR EMBL, AB024541; BAA89290.1; JOINED.  
 CC DR EMBL, AB024542; BAA89290.1; JOINED.  
 CC DR EMBL, AB024543; BAA89290.1; JOINED.  
 CC DR EMBL, AB024544; BAA89290.1; JOINED.  
 CC DR EMBL, AB024545; BAA89290.1; JOINED.  
 CC DR EMBL, AB024546; BAA89290.1; JOINED.  
 CC DR EMBL, AB024547; BAA89290.1; JOINED.  
 CC DR EMBL, AB024548; BAA89290.1; JOINED.  
 CC DR EMBL, AB024549; BAA89290.1; JOINED.  
 CC DR EMBL, AB024550; BAA89290.1; JOINED.  
 CC DR EMBL, AB024551; BAA89290.1; JOINED.  
 CC DR EMBL, AB024552; BAA89290.1; JOINED.  
 CC DR EMBL, AB024553; BAA89290.1; JOINED.  
 CC DR EMBL, AB024554; BAA89290.1; JOINED.  
 CC DR EMBL, AB024555; BAA89290.1; JOINED.  
 CC DR EMBL, AB024556; BAA89290.1; JOINED.  
 CC DR EMBL, AB024557; BAA89290.1; JOINED.  
 CC DR EMBL, AB024558; BAA89290.1; JOINED.  
 CC DR EMBL, AB024559; BAA89290.1; JOINED.  
 CC DR EMBL, AB024560; BAA89290.1; JOINED.  
 CC DR EMBL, AB024561; BAA89290.1; JOINED.  
 CC DR EMBL, AB024562; BAA89290.1; JOINED.  
 CC DR EMBL, AB024563; BAA89290.1; JOINED.  
 CC DR EMBL, AB024564; BAA89290.1; JOINED.  
 CC DR EMBL, AB024565; BAA89290.1; JOINED.  
 CC DR EMBL, AB024566; BAA89290.1; JOINED.  
 CC DR EMBL, AB024567; BAA89290.1; JOINED.  
 CC DR EMBL, AB024568; BAA89290.1; JOINED.  
 CC DR EMBL, AB024569; BAA89290.1; JOINED.  
 CC DR EMBL, AB024570; BAA89290.1; JOINED.  
 CC DR EMBL, AB024571; BAA89290.1; JOINED.  
 CC DR EMBL, AB024572; BAA89290.1; JOINED.  
 CC DR EMBL, AB024573; BAA89290.1; JOINED.  
 CC DR EMBL, AB024574; BAA89290.1; JOINED.  
 CC DR EMBL, AB024575; BAA89290.1; JOINED.  
 CC DR EMBL, AB024576; BAA89290.1; JOINED.  
 CC DR EMBL, AB024577; BAA89290.1; JOINED.  
 CC DR EMBL, AB024578; BAA89290.1; JOINED.  
 CC DR EMBL, AB024579; BAA89290.1; JOINED.  
 CC DR EMBL, AB024580; BAA89290.1; JOINED.  
 CC DR EMBL, AB024581; BAA89290.1; JOINED.  
 CC DR EMBL, AB024582; BAA89290.1; JOINED.  
 CC DR EMBL, AB024583; BAA89290.1; JOINED.  
 CC DR EMBL, AB024584; BAA89290.1; JOINED.  
 CC DR EMBL, AB024585; BAA89290.1; JOINED.  
 CC DR EMBL, AB024586; BAA89290.1; JOINED.  
 CC DR EMBL, AB024587; BAA89290.1; JOINED.  
 CC DR EMBL, AB024588; BAA89290.1; JOINED.  
 CC DR EMBL, AB024589; BAA89290.1; JOINED.  
 CC DR EMBL, AB024590; BAA89290.1; JOINED.  
 CC DR EMBL, AB024591; BAA89290.1; JOINED.  
 CC DR EMBL, AB024592; BAA89290.1; JOINED.  
 CC DR EMBL, AB024593; BAA89290.1; JOINED.  
 CC DR EMBL, AB024594; BAA89290.1; JOINED.  
 CC DR EMBL, AB024595; BAA89290.1; JOINED.  
 CC DR EMBL, AB024596; BAA89290.1; JOINED.  
 CC DR EMBL, AB024597; BAA89290.1; JOINED.  
 CC DR EMBL, AB024598; BAA89290.1; JOINED.  
 CC DR EMBL, AB024599; BAA89290.1; JOINED.  
 CC DR EMBL, AB024600; BAA89290.1; JOINED.  
 CC DR EMBL, AB024601; BAA89290.1; JOINED.  
 CC DR EMBL, AB024602; BAA89290.1; JOINED.  
 CC DR EMBL, AB024603; BAA89290.1; JOINED.  
 CC DR EMBL, AB024604; BAA89290.1; JOINED.  
 CC DR EMBL, AB024605; BAA89290.1; JOINED.  
 CC DR EMBL, AB024606; BAA89290.1; JOINED.  
 CC DR EMBL, AB024607; BAA89290.1; JOINED.  
 CC DR EMBL, AB024608; BAA89290.1; JOINED.  
 CC DR EMBL, AB024609; BAA89290.1; JOINED.  
 CC DR EMBL, AB024610; BAA89290.1; JOINED.  
 CC DR EMBL, AB024611; BAA89290.1; JOINED.  
 CC DR EMBL, AB024612; BAA89290.1; JOINED.  
 CC DR EMBL, AB024613; BAA89290.1; JOINED.  
 CC DR EMBL, AB024614; BAA89290.1; JOINED.  
 CC DR EMBL, AB024615; BAA89290.1; JOINED.  
 CC DR EMBL, AB024616; BAA89290.1; JOINED.  
 CC DR EMBL, AB024617; BAA89290.1; JOINED.  
 CC DR EMBL, AB024618; BAA89290.1; JOINED.  
 CC DR EMBL, AB024619; BAA89290.1; JOINED.  
 CC DR EMBL, AB024620; BAA89290.1; JOINED.  
 CC DR EMBL, AB024621; BAA89290.1; JOINED.  
 CC DR EMBL, AB024622; BAA89290.1; JOINED.  
 CC DR EMBL, AB024623; BAA89290.1; JOINED.  
 CC DR EMBL, AB024624; BAA89290.1; JOINED.  
 CC DR EMBL, AB024625; BAA89290.1; JOINED.  
 CC DR EMBL, AB024626; BAA89290.1; JOINED.  
 CC DR EMBL, AB024627; BAA89290.1; JOINED.  
 CC DR EMBL, AB024628; BAA89290.1; JOINED.  
 CC DR EMBL, AB024629; BAA89290.1; JOINED.  
 CC DR EMBL, AB024630; BAA89290.1; JOINED.  
 CC DR EMBL, AB024631; BAA89290.1; JOINED.  
 CC DR EMBL, AB024632; BAA89290.1; JOINED.  
 CC DR EMBL, AB024633; BAA89290.1; JOINED.  
 CC DR EMBL, AB024634; BAA89290.1; JOINED.  
 CC DR EMBL, AB024635; BAA89290.1; JOINED.  
 CC DR EMBL, AB024636; BAA89290.1; JOINED.  
 CC DR EMBL, AB024637; BAA89290.1; JOINED.  
 CC DR EMBL, AB024638; BAA89290.1; JOINED.  
 CC DR EMBL, AB024639; BAA89290.1; JOINED.  
 CC DR EMBL, AB024640; BAA89290.1; JOINED.  
 CC DR EMBL, AB024641; BAA89290.1; JOINED.  
 CC DR EMBL, AB024642; BAA89290.1; JOINED.  
 CC DR EMBL, AB024643; BAA89290.1; JOINED.  
 CC DR EMBL, AB024644; BAA89290.1; JOINED.  
 CC DR EMBL, AB024645; BAA89290.1; JOINED.  
 CC DR EMBL, AB024646; BAA89290.1; JOINED.  
 CC DR EMBL, AB024647; BAA89290.1; JOINED.  
 CC DR EMBL, AB024648; BAA89290.1; JOINED.  
 CC DR EMBL, AB024649; BAA89290.1; JOINED.  
 CC DR EMBL, AB024650; BAA89290.1; JOINED.  
 CC DR EMBL, AB024651; BAA89290.1; JOINED.  
 CC DR EMBL, AB024652; BAA89290.1; JOINED.  
 CC DR EMBL, AB024653; BAA89290.1; JOINED.  
 CC DR EMBL, AB024654; BAA89290.1; JOINED.  
 CC DR EMBL, AB024655; BAA89290.1; JOINED.  
 CC DR EMBL, AB024656; BAA89290.1; JOINED.  
 CC DR EMBL, AB024657; BAA89290.1; JOINED.  
 CC DR EMBL, AB024658; BAA89290.1; JOINED.  
 CC DR EMBL, AB024659; BAA89290.1; JOINED.  
 CC DR EMBL, AB024660; BAA89290.1; JOINED.  
 CC DR EMBL, AB024661; BAA89290.1; JOINED.  
 CC DR EMBL, AB024662; BAA89290.1; JOINED.  
 CC DR EMBL, AB024663; BAA89290.1; JOINED.  
 CC DR EMBL, AB024664; BAA89290.1; JOINED.  
 CC DR EMBL, AB024665; BAA89290.1; JOINED.  
 CC DR EMBL, AB024666; BAA89290.1; JOINED.  
 CC DR EMBL, AB024667; BAA89290.1; JOINED.  
 CC DR EMBL, AB024668; BAA89290.1; JOINED.  
 CC DR EMBL, AB024669; BAA89290.1; JOINED.  
 CC DR EMBL, AB024670; BAA89290.1; JOINED.  
 CC DR EMBL, AB024671; BAA89290.1; JOINED.  
 CC DR EMBL, AB024672; BAA89290.1; JOINED.  
 CC DR EMBL, AB024673; BAA89290.1; JOINED.  
 CC DR EMBL, AB024674; BAA89290.1; JOINED.  
 CC DR EMBL, AB024675; BAA89290.1; JOINED.  
 CC DR EMBL, AB024676; BAA89290.1; JOINED.  
 CC DR EMBL, AB024677; BAA89290.1; JOINED.  
 CC DR EMBL, AB024678; BAA89290.1; JOINED.  
 CC DR EMBL, AB024679; BAA89290.1; JOINED.  
 CC DR EMBL, AB024680; BAA89290.1; JOINED.  
 CC DR EMBL, AB024681; BAA89290.1; JOINED.  
 CC DR EMBL, AB024682; BAA89290.1; JOINED.  
 CC DR EMBL, AB024683; BAA89290.1; JOINED.  
 CC DR EMBL, AB024684; BAA89290.1; JOINED.  
 CC DR EMBL, AB024685; BAA89290.1; JOINED.  
 CC DR EMBL, AB024686; BAA89290.1; JOINED.  
 CC DR EMBL, AB024687; BAA89290.1; JOINED.  
 CC DR EMBL, AB024688; BAA89290.1; JOINED.  
 CC DR EMBL, AB024689; BAA89290.1; JOINED.  
 CC DR EMBL, AB024690; BAA89290.1; JOINED.  
 CC DR EMBL, AB024691; BAA89290.1; JOINED.  
 CC DR EMBL, AB024692; BAA89290.1; JOINED.  
 CC DR EMBL, AB024693; BAA89290.1; JOINED.  
 CC DR EMBL, AB024694; BAA89290.1; JOINED.  
 CC DR EMBL, AB024695; BAA89290.1; JOINED.  
 CC DR EMBL, AB024696; BAA89290.1; JOINED.  
 CC DR EMBL, AB024697; BAA89290.1; JOINED.  
 CC DR EMBL, AB024698; BAA89290.1; JOINED.  
 CC DR EMBL, AB024699; BAA89290.1; JOINED.  
 CC DR EMBL, AB024700; BAA89290.1; JOINED.  
 CC DR EMBL, AB024701; BAA89290.1; JOINED.  
 CC DR EMBL, AB024702; BAA89290.1; JOINED.  
 CC DR EMBL, AB024703; BAA89290.1; JOINED.  
 CC DR EMBL, AB024704; BAA89290.1; JOINED.  
 CC DR EMBL, AB024705; BAA89290.1; JOINED.  
 CC DR EMBL, AB024706; BAA89290.1; JOINED.  
 CC DR EMBL, AB024707; BAA89290.1; JOINED.  
 CC DR EMBL, AB024708; BAA89290.1; JOINED.  
 CC DR EMBL, AB024709; BAA89290.1; JOINED.  
 CC DR EMBL, AB024710; BAA89290.1; JOINED.  
 CC DR EMBL, AB024711; BAA89290.1; JOINED.  
 CC DR EMBL, AB024712; BAA89290.1; JOINED.  
 CC DR EMBL, AB024713; BAA89290.1; JOINED.  
 CC DR EMBL, AB024714; BAA89290.1; JOINED.  
 CC DR EMBL, AB024715; BAA89290.1; JOINED.  
 CC DR EMBL, AB024716; BAA89290.1; JOINED.  
 CC DR EMBL, AB024717; BAA89290.1; JOINED.  
 CC DR EMBL, AB024718; BAA89290.1; JOINED.  
 CC DR EMBL, AB024719; BAA89290.1; JOINED.  
 CC DR EMBL, AB024720; BAA89290.1; JOINED.  
 CC DR EMBL, AB024721; BAA89290.1; JOINED.  
 CC DR EMBL, AB024722; BAA89290.1; JOINED.  
 CC DR EMBL, AB024723; BAA89290.1; JOINED.  
 CC DR EMBL, AB024724; BAA89290.1; JOINED.  
 CC DR EMBL, AB024725; BAA89290.1; JOINED.  
 CC DR EMBL, AB024726; BAA89290.1; JOINED.  
 CC DR EMBL, AB024727; BAA89290.1; JOINED.  
 CC DR EMBL, AB024728; BAA89290.1; JOINED.  
 CC DR EMBL, AB024729; BAA89290.1; JOINED.  
 CC DR EMBL, AB024730; BAA89290.1; JOINED.  
 CC DR EMBL, AB024731; BAA89290.1; JOINED.  
 CC DR EMBL, AB024732; BAA89290.1; JOINED.  
 CC DR EMBL, AB024733; BAA89290.1; JOINED.  
 CC DR EMBL, AB024734; BAA89290.1; JOINED.  
 CC DR EMBL, AB024735; BAA89290.1; JOINED.  
 CC DR EMBL, AB024736; BAA89290.1; JOINED.  
 CC DR EMBL, AB024737; BAA89290.1; JOINED.  
 CC DR EMBL, AB024738; BAA89290.1; JOINED.  
 CC DR EMBL, AB024739; BAA89290.1; JOINED.  
 CC DR EMBL, AB024740; BAA89290.1; JOINED.  
 CC DR EMBL, AB024741; BAA89290.1; JOINED.  
 CC DR EMBL, AB024742; BAA89290.1; JOINED.  
 CC DR EMBL, AB024743; BAA89290.1; JOINED.  
 CC DR EMBL, AB024744; BAA89290.1; JOINED.  
 CC DR EMBL, AB024745; BAA89290.1; JOINED.  
 CC DR EMBL, AB024746; BAA89290.1; JOINED.  
 CC DR EMBL, AB024747; BAA89290.1; JOINED.  
 CC DR EMBL, AB024748; BAA89290.1; JOINED.  
 CC DR EMBL, AB024749; BAA89290.1; JOINED.  
 CC DR EMBL, AB024750; BAA89290.1; JOINED.  
 CC DR EMBL, AB024751; BAA89290.1; JOINED.  
 CC DR EMBL, AB024752; BAA89290.1; JOINED.  
 CC DR EMBL, AB024753; BAA89290.1; JOINED.  
 CC DR EMBL, AB024754; BAA89290.1; JOINED.  
 CC DR EMBL, AB024755; BAA89290.1; JOINED.  
 CC DR EMBL, AB024756; BAA89290.1; JOINED.  
 CC DR EMBL, AB024757; BAA89290.1; JOINED.  
 CC DR EMBL, AB024758; BAA89290.1; JOINED.  
 CC DR EMBL, AB024759; BAA89290.1; JOINED.  
 CC DR EMBL, AB024760; BAA89290.1; JOINED.  
 CC DR EMBL, AB024761; BAA89290.1; JOINED.  
 CC DR EMBL, AB024762; BAA89290.1; JOINED.  
 CC DR EMBL, AB024763; BAA89290.1; JOINED.  
 CC DR EMBL, AB024764; BAA89290.1; JOINED.  
 CC DR EMBL, AB024765; BAA89290.1; JOINED.  
 CC DR EMBL, AB024766; BAA89290.1; JOINED.  
 CC DR EMBL, AB024767; BAA89290.1; JOINED.  
 CC DR EMBL, AB024768; BAA89290.1; JOINED.  
 CC DR EMBL, AB024769; BAA89290.1; JOINED.  
 CC DR EMBL, AB024770; BAA89290.1; JOINED.  
 CC DR EMBL, AB024771; BAA89290.1; JOINED.  
 CC DR EMBL, AB024772; BAA89290.1; JOINED.  
 CC DR EMBL, AB024773; BAA89290.1; JOINED.  
 CC DR EMBL, AB024774; BAA89290.1; JOINED.  
 CC DR EMBL, AB024775; BAA89290.1; JOINED.  
 CC DR EMBL, AB024776; BAA89290.1; JOINED.  
 CC DR EMBL, AB024777; BAA89290.1; JOINED.  
 CC DR EMBL, AB024778; BAA89290.1; JOINED.  
 CC DR EMBL, AB024779; BAA89290.1; JOINED.  
 CC DR EMBL, AB024780; BAA89290.1; JOINED.  
 CC DR EMBL, AB024781; BAA89290.1; JOINED.  
 CC DR EMBL, AB024782; BAA89290.1; JOINED.  
 CC DR EMBL, AB024783; BAA89290.1; JOINED.  
 CC DR EMBL, AB024784; BAA89290.1; JOINED.  
 CC DR EMBL, AB024785; BAA89290.1; JOINED.  
 CC DR EMBL, AB024786; BAA89290.1; JOINED.  
 CC DR EMBL, AB024787; BAA89290.1; JOINED.  
 CC DR EMBL, AB024788; BAA89290.1; JOINED.  
 CC DR EMBL, AB024789; BAA89290.1; JOINED.  
 CC DR EMBL, AB024790; BAA89290.1; JOINED.  
 CC DR EMBL, AB024791; BAA89290.1; JOINED.  
 CC DR EMBL, AB024792; BAA89290.1; JOINED.  
 CC DR EMBL, AB024793; BAA89290.1; JOINED.  
 CC DR EMBL, AB024794; BAA89290.1; JOINED.  
 CC DR EMBL, AB024795; BAA89290.1; JOINED.  
 CC DR EMBL, AB024796; BAA89290.1; JOINED.  
 CC DR EMBL, AB024797; BAA89290.1; JOINED.  
 CC DR EMBL, AB024798; BAA89290.1; JOINED.  
 CC DR EMBL, AB024799; BAA89290.1; JOINED.  
 CC DR EMBL, AB024800; BAA89290.1; JOINED.  
 CC DR EMBL, AB024801; BAA89290.1; JOINED.  
 CC DR EMBL, AB024802; BAA89290.1; JOINED.  
 CC DR EMBL, AB024803; BAA89290.1; JOINED.  
 CC DR EMBL, AB024804; BAA89290.1; JOINED.  
 CC DR EMBL, AB024805; BAA89290.1; JOINED.  
 CC DR EMBL, AB024806; BAA89290.1; JOINED.  
 CC DR EMBL, AB024807; BAA89290.1; JOINED.  
 CC DR EMBL, AB024808; BAA89290.1; JOINED.  
 CC DR EMBL, AB024809; BAA89290.1; JOINED.  
 CC DR EMBL, AB024810; BAA89290.1; JOINED.  
 CC DR EMBL, AB024811; BAA89290.1; JOINED.  
 CC DR EMBL, AB024812; BAA89290.1; JOINED.  
 CC DR EMBL, AB024813; BAA89290.1; JOINED.  
 CC DR EMBL, AB024814; BAA89290.1; JOINED.  
 CC DR EMBL, AB024815; BAA89290.1; JOINED.  
 CC DR EMBL, AB024816; BAA89290.1; JOINED.  
 CC DR EMBL, AB024817; BAA89290.1; JOINED.  
 CC DR EMBL, AB024818; BAA89290.1; JOINED.  
 CC DR EMBL, AB024819; BAA89290.1; JOINED.  
 CC DR EMBL, AB024820; BAA89290.1; JOINED.  
 CC DR EMBL, AB024821; BAA89290.1; JOINED.  
 CC DR EMBL, AB024822; BAA89290.1; JOINED.  
 CC DR EMBL, AB024823; BAA89290.1; JOINED.  
 CC DR EMBL, AB024824; BAA89290.1; JOINED.  
 CC DR EMBL, AB024825; BAA89290.1; JOINED.  
 CC DR EMBL, AB024826; BAA89290.1; JOINED.  
 CC DR EMBL, AB024827; BAA89290.1; JOINED.  
 CC DR EMBL, AB024828; BAA89290.1; JOINED.  
 CC DR EMBL, AB024829; BAA89290.1; JOINED.  
 CC DR EMBL, AB024830; BAA89290.1; JOINED.  
 CC DR EMBL, AB024831; BAA89290.1; JOINED.  
 CC DR EMBL, AB024832; BAA89290.1; JOINED.  
 CC DR EMBL, AB024833; BAA89290.1; JOINED.  
 CC DR EMBL, AB024834; BAA89290.1; JOINED.  
 CC DR EMBL, AB024835; BAA89290.1; JOINED.  
 CC DR EMBL, AB024836; BAA89290.1; JOINED.  
 CC DR EMBL, AB024837; BAA89290.1; JOINED.  
 CC DR EMBL, AB024838; BAA89290.1; JOINED.  
 CC DR EMBL, AB024839; BAA89290.1; JOINED.  
 CC DR EMBL, AB024840; BAA89290.1; JOINED.  
 CC DR EMBL, AB024841; BAA89290.1; JOINED.  
 CC DR EMBL, AB024842; BAA89290.1; JOINED.  
 CC DR EMBL, AB024843; BAA89290.1; JOINED.  
 CC DR EMBL, AB024844; BAA89290.1; JOINED.  
 CC DR EMBL, AB024845; BAA89290.1; JOINED.  
 CC DR EMBL, AB024846; BAA89290.1; JOINED.  
 CC DR EMBL, AB024847; BAA89290.1; JOINED.  
 CC DR EMBL, AB024848; BAA89290.1; JOINED.  
 CC DR EMBL, AB024849; BAA89290.1; JOINED.  
 CC DR EMBL, AB024850; BAA89290.1; JOINED.  
 CC DR EMBL, AB024851; BAA89290.1; JOINED.  
 CC DR EMBL, AB024852; BAA89290.1; JOINED.  
 CC DR EMBL, AB024853; BAA89290.1; JOINED.  
 CC DR EMBL, AB024854; BAA89290.1; JOINED.  
 CC DR EMB

RP 446-453, N-GLYCOSYLATION, PHOSPHORYLATION ON TYROSINE RESIDUES, AND  
 RP INTERACTIONS WITH PTPN6 AND PTPN11.  
 RC TISSUE=Fetal fibroblast;  
 RX MEDLINE=97096667; PubMed=8943344;  
 RA Fujioke Y., Matczak T., Noguchi T., Iwamatsu A., Yamao T.,  
 RA Takahashi N., Tauda M., Takeda T., Kaunga M.;  
 RA "A novel membrane glycoprotein, SHS-1, that binds the SH2 domain-  
 RT containing protein tyrosine phosphatase SHP-2 in response to mitogens  
 RT and cell adhesion.";  
 RL Mol. Cell. Biol. 16:6887-6893(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 32-48 AND 446-453, FUNCTION, AND  
 RP PHOSPHORYLATION ON TYROSINE RESIDUES.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=97415431; PubMed=92712230;  
 RA Sano S.-I., Ohnishi H., Omori A., Hasegawa J., Kubota M.;  
 RA "B1T, an immune antigen receptor A-like molecule in the brain.";  
 RL FEBS Lett. 411:327-334(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 99-107, 128-149, 192-217, 405-417;  
 RP 419-429; 446-467 AND 496-506, N-GLYCOSYLATION, AND TISSUE SPECIFICITY.  
 RC STRAIN=Fischer 344; TISSUE=Skin;  
 RX MEDLINE=96449911; PubMed=9774638;  
 RA Segnario C., Sterling H., Beckers C., Kobayashi R., Solimena M.,  
 RA Ullis E., Vignery A.;  
 RT "MR, a putative receptor mediates the fusion of macrophages.";  
 RL Mol. Cell. Biol. 18:6213-6223(1998).  
 RN [4]  
 RP SEQUENCE OF 1-419 FROM N.A., AND TISSUE SPECIFICITY.  
 RC STRAIN=MAG/Rij; TISSUE=Alveolar macrophage;  
 RX MEDLINE=96375871; PubMed=9712053;  
 RA Adams S., van der Laan L.J.W., Verdon-Wilson E.,  
 RA Renardel de Lavalette C., Doepf E.A., Dijkstra C.D., Simmonds D.L.,  
 RA van den Berg T.K.;  
 RT "Signal-regulatory protein is selectively expressed by myeloid and  
 RT neuronal cells.";  
 RL J. Immunol. 161:1853-1859(1998).  
 RN [5]  
 RP PHOSPHORYLATION IN RESPONSE TO EGF, AND INTERACTION WITH PTPN11.  
 RX MEDLINE=9800885; PubMed=9344856;  
 RA Ochi F., Matczak T., Noguchi T., Fujioke Y., Yamao T., Takeda T.,  
 RA Tauda M., Takeda H., Fukunaga K., Okabayashi Y., Kaunga M.;  
 RT "Epidermal growth factor stimulates the tyrosine phosphorylation of  
 RT SHS-1 and association of SHS-1 with SHP-2, a SH2 domain-containing  
 RT protein tyrosine phosphatase.";  
 RL Biochem. Biophys. Res. Commun. 239:483-487(1997).  
 RN [6]  
 RP PHOSPHORYLATION OF TYR-477 AND TYR-501, AND MUTAGENESIS OF TYR-436;  
 RP TYR-460; TYR-477 AND TYR-501.  
 RX MEDLINE=98204923; PubMed=9535915;  
 RA Takeda T., Matczak T., Takeda H., Fukunaga K., Noguchi T.,  
 RA Fujioke Y., Okazaki I., Tauda M., Yamao T., Ochi F., Kaunga M.;  
 RT "Role of the complex formation of SHS-1 with SHP-2 in  
 RT insulin-stimulated mitogen-activated protein kinase activation.";  
 RL J. Biol. Chem. 273:9234-9242(1998).  
 RN [7]  
 RP FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts  
 CC as docking protein and induces translocation of PRP6,  
 CC PTPN11 and other binding partners from the cytosol to the  
 CC plasma membrane. Supports adhesion of cerebellar neurons, neurite  
 CC outgrowth and glial cell attachment. May play a key role in  
 CC intracellular signaling during synaptogenesis and in synaptic  
 CC function. Involved in the negative regulation of receptor tyrosine  
 CC kinase-coupled cellular responses induced by cell adhesion, growth  
 CC factors or insulin. Mediates negative regulation of phagocytosis,  
 CC mast cell activation and dendritic cell activation. CD47 binding  
 CC prevents maturation of immature dendritic cells and inhibits  
 CC cytokine production by mature dendritic cells. May play a role in  
 CC the release of nitric oxide by macrophages (By similarity).  
 CC -1 SUBUNIT: Binds PTPN11 when tyrosine phosphorylated, except in  
 CC macrophages, where it primarily binds PRP6. Binds GNB2 in vitro.  
 CC Binds PdgR. Binds JAK2 irrespective of its phosphorylation status  
 CC and forms a stable complex. Binds SCAPI and/or SCAPI2. The  
 CC resulting complex recruits Fyb. Binds PRX28 (By similarity).

CC	-1- SURCELLULAR LOCATION: Type I membrane protein.
CC	-1- TISSUE SPECIFICITY: Highly expressed in brain, spleen, lung, liver
CC	and kidney. Detected at lower levels in heart. Highly expressed in
CC	alveolar and peritoneal macrophages, and at lower levels in
CC	dendritic cells.
CC	-1- PTM: N-glycosylated.
CC	-1- PM: Phosphorylated on tyrosine residues in response to insulin,
CC	cell adhesion or epidermal growth factors. Dephosphorylated by
CC	pTEN1.
CC	-1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
CC	-1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC	-- --
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isdb.ch/announce/
CC	or send an email to license@isdb-sib.ch).
CC	-- --
DR	EMBL; D86183; BAA12734.1; -.
DR	EMBL; U62328; AACG9478.1; -.
DR	EMBL; AF055065; AACC18089.1; -.
DR	HSSP; P01703; FPAB.
DR	InterPro; IPRO07110; IG_1like.
DR	InterPro; IPRO03597; IG_C1.
DR	InterPro; IPRO03006; IG_MHC. Pfam; PF00047; Ig_f3.
DR	SMART; SMO0407; IgC1; 2.
DR	PROSITE; PS50835; IG_LIKE; 3.
DR	PROSITE; PS00290; IG_MHC; 1.
KW	Repeat; Signal; Transmembrane; Immunoglobulin domain; SH3-binding;
KV	Glycoprotein; Phosphorylation.
FT	SIGNAL 1 31
FT	CCHAIN 32 509
FT	TYPE SUBSTRATE 1.
FT	DOMAIN 28 373
FT	TRANSMEM 374 394
FT	POTENTIAL.
FT	DOMAIN 395 509
FT	CYTOPLASMIC (POTENTIAL).
FT	DONAIN 32 138
FT	IG-LIKE V-TYPE.
FT	DONAIN 150 248
FT	IG-LIKE C1-TYPE 1.
FT	DONAIN 255 349
FT	IG-LIKE C1-TYPE 2.
FT	DISULFID 55 122 POTENTIAL.
FT	DISULFID 172 229 POTENTIAL.
FT	DISULFID 274 332 POTENTIAL.
FT	SITE 436 439
FT	SITE 446 451
FT	SITE 460 463
FT	SITE 477 480
FT	SITE 501 504
FT	MOD_RES 436 436
FT	(PHOSPHORYLATION (BY TYR-KINASES)
FT	(POTENTIAL)).
FT	MOD_RES 460 460
FT	(PHOSPHORYLATION (BY TYR-KINASES)
FT	(POTENTIAL)).
MOD_RES 477 477	PHOSPHORYLATION (BY TYR-KINASES) .
F MOD_RES 501 501	PHOSPHORYLATION (BY TYR-KINASES) .
FT CARBOHYD 54 54	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 93 93	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 169 169	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 181 181	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 205 205	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 209 209	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 242 242	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 246 246	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 271 271	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 293 293	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 312 312	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 320 320	N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 345 345	N-LINKED (GLCNAC... ) (POTENTIAL).
FT MUTAGEN 436 436	Y->: ABOLISHES TYROSINE PHOSPHORYLATION AND PTEN1 BINDING; WHEN ASSOCIATED WITH F-460 F-477 AND F-501.

```
FT MUTAGEN 460 460 Y->F: ABOLISHES TYROSINE PHOSPHORYLATION
FT PTPN11 BINDING, WHEN ASSOCIATED WITH
FT F-436, F-477 AND F-501.
FT Y->F: STRONGLY REDUCES INSULIN-INDUCED
FT TYROSINE PHOSPHORYLATION AND PTPN11
FT BINDING. ABOLISHES TYROSINE
FT PHOSPHORYLATION AND PTPN11 BINDING, WHEN
FT ASSOCIATED WITH F-436, F-460 AND F-501.
FT Y->F: STRONGLY REDUCES INSULIN-INDUCED
FT TYROSINE PHOSPHORYLATION AND PTPN11
FT BINDING. ABOLISHES TYROSINE
FT PHOSPHORYLATION AND PTPN11 BINDING, WHEN
FT ASSOCIATED WITH F-436, F-460 AND F-477.
FT MUTAGEN 501 501
FT
FT
FT CONFLICT 8 8
FT CONFLICT 10 10
FT CONFLICT 25 25
FT CONFLICT 58 58
FT CONFLICT 99 100
FT CONFLICT 162 162
FT CONFLICT 189 189
FT CONFLICT 205 205
FT CONFLICT 209 209
FT CONFLICT 405 405
FT CONFLICT 416 416
FT CONFLICT 418 421
FT CONFLICT 450 450
FT CONFLICT 499 499
SQ SEQUENCE 509 AA, 55690 MW, 581F60A4DD429F4 CRC64,
Query Match 7.1%; Score 192; DB 1; Length 509;
Best Local Similarity 21.2%; Pred. No. 3,8e-05;
Matches 105; Conservative 74; Mismatches 159; Indels 158; Gaps 21;
23 TCGNVLKGGDYELTCTAQSISI-----QFHKNKSNQITLNGQSFLTKSGSKND 78
37 TQADSVSVAAADSLTCTVSSLPVPGIKWKGEGQRSEIYFSGEHPRTITVSD 96
79 RADSRSLWDQGNPFLIKLKIEDSDTYICEVDQKEVQLVGLTANSPHLLQOS 138
97 -ATKNNM-----DFICISNVPEDAGTYC-VKQKQGV-----PDETKSGGG 141
139 LTTLESPPGSSPVQCSPPGKNIQGGKTLVSQLELDQSGTWTCTVNLQNKVEFKID 198
142 TLLVYLAKP-SSPEVSGPDSRG---SPGQTVN-----FTC----- 172
199 IVPCAPEKSCDKHTHCCELLGSPSVLPFPKPKDITLMIKRTPEVTCVVDVSHEDPEV 258
173 -----KS-----YGFSPR-----NI 182
259 KENMYVGVENNAKTKPREQYNSTYRVSVLTV-LHQDMLNGEKYCKKYNKALPA-P 316
183 TLKMLKDKGKELSHLETTSKKS-NVSYNISTSVSKLSPEDHSHVIT-CEVAHVLTBERP 240
317 IEKTSKAKGQPREQVYTLTPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGPENN 375
241 INGTANFSLIIRVSP---TLKITQQLTPASQVNLTCQVKPKRALQLNMLENGNLST 297
376 YKTPPVLDSSGFFLYSKLTVDKSRMKGAGNFGSSVWHE---AAHNYTQKSLSLSGCL 432
298 DPEHFTDNRDGTNYVTSLEFLVNSAHREDVFTQVHDSQPALTEHTHYAFAFHS--- 354
433 QLDFTCAENQDELGLGLWTTDPPRASALPAPETGSAALPDQPSAALPDPAASALPALA 492
355 -----SSGGSME-----TIFDNN-----A 368
493 VISFLGLGLGVACVL 508
369 YNNMNVFVIGVACAL 384
RESULT 61
SRB2_HUMAN STANDARD; PRT; 387 AA.
```

```
AC Q9P1W8; Q8WMA5; Q9NQK8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Signal-regulatory protein beta-2 precursor (SIRP-beta-2) (SIRP-b2).
GN SIRPB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Placenta;
RX MEDLINE=2103165; PubMed=11165750;
RA Ichigotani Y., Matsuda S., Machida K., Oshima K., Iwamoto T.,
RT Yamaki K., Hayakawa T., Hamaguchi M.,
RT "Molecular cloning of a novel human gene (SIRP-B2) which encodes a new
RT member of the SIRP/SHPS-1 protein family.";
RL J. Hum. Genet. 45:378-382(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638745; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ahnurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Codley V.E., Collier R.E., Connor R.E., Cobby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Griffiths D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharvalho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., Murray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,
RA Rice C.M., Ross M.T., Scott C.E., Senta H.K., Showkseen R., Sims S.,
RA Skuse C.D., Smith M.L., Soderlund C., Steward G.A., Sultson J.E.,
RA Swann R.M., Sycamore N., Taylor R., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Trean A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whithead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Lung;
RX MEDLINE=22386257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shuman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stiepleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohshiyuki S., Carcinici P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grilwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.U.M., Matra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```
CC -1- FUNCTION: Probable immunoglobulin-like cell surface receptor.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9P1W8-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9P1W8-2; Sequence=VSP_007027;
CC Note=No experimental confirmation available;
CC Name=3;
CC IsoId=Q9P1W8-3; Sequence=VSP_007028;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Detected in liver, and at very low levels in
CC brain, heart, lung, pancreas, kidney, placenta and skeletal
CC muscle.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AB042624; BAA95692.1; -.
CC EMBL, AL138804; CAC00474.1; -.
CC EMBL, BC020629; AAH20629.1; ALT_INIT.
CC GenBank, HGNC:15757; SIRPB2.
CC MIM, 605466; -.
CC DR GO, GO:0007267; P:cell-cell signaling; TAS.
CC DR GO, GO:0007242; P:intracellular signaling cascade; TAS.
CC DR GO, GO:0008285; P:negative regulation of cell proliferation; TAS.
CC DR InterPro, IPR007110; IG_1like.
CC DR InterPro, IPR003597; IG_c1.
CC DR InterPro, IPR003006; IG_MHC.
CC DR InterPro, IPR003596; IG_v.
CC DR Pfam, PF00047; Ig_3.
CC DR SMART, SM00407; IgC1; 2.
CC DR SMART, SM00406; IgV; 1.
CC DR PROSITE, PS00835; IG_LIKE; 3.
CC DR PROSITE, PS00290; IG_MHC; FALSE NEG.
CC KM Repeat; Signal; Transmembrane; Immunoglobulin domain; Glycoprotein;
CC Alternative splicing.
CC FT SIGNAL 1 28 POTENTIAL.
CC FT CHAIN 29 387 SIGNAL-REGULATORY PROTEIN BETA-2.
CC FT DOMAIN 29 360 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 361 383 POTENTIAL.
CC FT DOMAIN 384 387 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 29 137 IG-LIKE V-TYPE.
CC FT DOMAIN 146 245 IG-LIKE C1-TYPE 1.
CC FT DOMAIN 252 340 IG-LIKE C1-TYPE 2.
CC FT DISULFID 53 119 POTENTIAL.
CC FT DISULFID 168 226 POTENTIAL.
CC FT DISULFID 221 328 POTENTIAL.
CC FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT VARSPLIC 1 33 Missing (in isoform 2).
CC FT VARSPLIC 33 /FridaVSP_007027.
CC FT VARSPLIC 144 360 Missing (in isoform 3).
CC FT VARSPLIC 144 /FridaVSP_007028.
CC FT VARSPLIC 263 263 A -> V (IN REF. 2).
CC FT VARSPLIC 266 266 L -> S (IN REF. 2).
CC FT VARSPLIC 387 387 A -> V (IN REF. 2).
CC SEQUENCE 387 AA; 42495 MW; F7P20C9P86805E4B CRC64;
CC Query Match 7.0%; Score 188.5; DB 1; Length 387;
CC Best Local Similarity 21.9%; Pred. No. 4.5e-05;
CC Matches 113; Conservative 56; Mismatches 174; Indels 173; Gaps 20;
```

```
CC Db 13 PF--LLTLTLTLQTEVAGEEELQMIOPEKLLLVYTGKATLACTYT----- 56
CC QY 59 IKILNQGSFLLTKGSPSKLNDRADSRRLMDQGNP--LIINKLKI ESDTYICEVEDOK 115
CC Db 57 -----SLLPVGP-----VLMFRGVGPRELITY-----NQK 81
CC QY 116 EKVQVLVRLTANSTHLLQGSLLVLTLESPPGSSPSVQCRRPRKNIQCKTLVSQLE 175
CC Db 82 E-----GHPRTVTVSDLTRNNMDFSRISIT 110
CC QY 176 LQDSGTWCTVLQ--NOKKVEKIDIVCPAPAPKSCDKTHCPGLGSPV-FLFPPKP 232
CC Db 111 PADVGTTCVKKRKSPENVEK-----SGPGTEMLGAKRP 146
CC QY 233 KQTLNIS--RPEVTCVVVDVSH--EDPEVKFMVYDGVFNHAKTKPREQYNSTYR 286
CC Db 147 SAPVILGPAARTTPRHVTSFTCSHGFSPRDITLKMFKNGNELSPQFVVDPTQSVAYS 206
CC QY 287 VVSULTVLHODPLNGEKYKCKSNKALRA-PIEKT--ISKAGQREPOVYTLPPSRDEL 343
CC Db 207 IKSTARVVLDPDVDSQVYCEVAHTVTLQDPLRGTRANSEAIRVPTLEV-TQQPMR--- 262
CC QY 344 TKNQVSLTCLVKGFPYSDIAVEMESNGQENNYKTPPYLDSGSPFLYSKLYDKSRWQ 403
CC Db 263 AGNQVAVTCQVRKFPQSLQTLTWLENGVCQRETSLTLENKQGYTNMTSPLVVISDQR 322
CC QY 404 QGNVSCSVMEBALNHNTOKSLSPGLQIDETCAEADQGLDGLMTTPDPPASALPAP 463
CC Db 323 DDVVLTLQCVKHG-----QLAVSKRLALEVTVHQXD----- 353
CC QY 464 PTGSALPDPQTSALPDPASALPALAVISFLIG 499
CC Db 354 -----QSSDATGPP--ASSL-TALLLAVLIG 377
CC RESULT 62
CC SHS1_BOVIN STANDARD; PRT; 506 AA.
CC ID SHS1_BOVIN
CC AC 046631; 046632;
CC DT 10-OCT-2003 (Rel. 42, Created)
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor
CC DE (SHP substrate-1) (SHPs-1) (Inhibitory receptor SHS-1) (Signal-
CC DE regulatory protein alpha-1) (S1rp-alpha-1) (MyD-1 antigen).
CC DE PTPNS1 OR SHPS1 OR S1RP OR MYD1.
CC OS Bos taurus (Bovine).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
CC OC Bovidae; Bovinae; Bos.
CC NCBI_TaxId=9913;
CC [1]
CC SEQUENCE FROM N.A., AND VARIANTS SER-23; ALA-28; LEU-61; ARG-70;
CC HIS-120; I25-GLN; I27-GLY; I29-HIS; I32-VAL; ASN-145; VAL-155;
CC ASP-203; ARG-261; LEU-302; LEU-316; ARG-337; ASN-367; LEU-422; PHE-429
CC RP AND GLU-433.
CC STRAIN=Frilesian; TISSUE=peripheral blood;
CC MEDLINE=98143722; PubMed=9485180;
CC Brooke G.P., Parsons K.R., Howard C.J.;
CC RA Cloning of two members of the S1RP alpha family of protein tyrosine
CC RT phosphatase binding proteins in cattle that are expressed on monocytes
CC RT and a subpopulation of dendritic cells and which mediate binding to
CC RT CD4 T cells."
CC RL Eur. J. Immunol. 28:1-11(1998).
CC CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts
CC as docking protein and induces translocation of PIP2,
CC PIP3 and other binding partners from the cytosol to the
CC plasma membrane. Supports adhesion of cerebellar neurons, neurite
CC outgrowth and glial cell attachment. May play a key role in
CC intracellular signaling during synaptogenesis and in synaptic
CC function. Involved in the negative regulation of
CC receptor tyrosine kinase-coupled cellular responses induced by
```



CC cell adhesion, growth factors or insulin. Mediates negative  
 CC regulation of phagocytosis, mast cell activation and dendritic  
 CC cell activation. CD47 binding prevents maturation of immature  
 CC dendritic cells and inhibits cytokine production by mature  
 CC dendritic cells (By similarity).  
 CC -1 SUBUNIT: Binds PRPN1 when tyrosine-phosphorylated, except in  
 CC macrophages, where it primarily binds PRPN6. Binds GIB2 in vitro.  
 CC Binds JAK2 irrespective of its phosphorylation status and forms a  
 CC stable complex. Binds SCAP1 and/or SCAP2. The resulting complex  
 CC recruits Fyb. Binds FGR and PTK2B (By similarity).  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1 TISSUE SPECIFICITY: Highly expressed in spleen macrophages.  
 CC -1 Detected in skin dendritic cells.  
 CC -1 PTM: Phosphorylated on tyrosine residues (By similarity).  
 CC -1 SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.  
 CC -1 SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL, Y11045; CA71942.1; -  
 CC DR EMBL, Y11046; CA71943.1; -  
 CC DR InterPro: IPR003597; IG-like.  
 CC DR InterPro: IPR003006; IG-cl.  
 CC DR Pfam; PF00047; IG; 3.  
 CC DR SMART; SM00407; IGcl; 2.  
 CC DR PROSITE; PS00835; IG\_LIKE; 3.  
 CC DR PROSITE; PS00290; IG\_MHC; FALSE NEG.  
 CC Repeat; Signal; Transmembrane; Immunoglobulin domain; SH3-binding;  
 CC Glycoprotein; Phosphorylation; Polymorphism.  
 CC KW SIGNAL 1 29  
 CC CHAIN 30 506  
 CC FT DOMAIN 30 371  
 CC FT TRANSMEM 372 392  
 CC FT DOMAIN 393 506  
 CC FT DOMAIN 30 145  
 CC FT DOMAIN 148 248  
 CC FT DOMAIN 255 348  
 CC FT DISULFID 55 121  
 CC FT DISULFID 170 228  
 CC FT DISULFID 273 331  
 CC FT SITE 432 435  
 CC FT SITE 441 446  
 CC FT SITE 455 458  
 CC FT SITE 472 475  
 CC FT SITE 498 501  
 CC FT MOD\_RES 431 431  
 CC FT MOD\_RES 455 455  
 CC FT MOD\_RES 472 472  
 CC FT MOD\_RES 498 498  
 CC FT CARBOHYD 92 92  
 CC FT CARBOHYD 167 167  
 CC FT CARBOHYD 179 179  
 CC FT CARBOHYD 204 204  
 CC FT CARBOHYD 210 210  
 CC FT CARBOHYD 246 246  
 CC FT CARBOHYD 270 270  
 CC FT CARBOHYD 292 292  
 CC FT CARBOHYD 311 311  
 CC FT CARBOHYD 319 319  
 CC FT CARBOHYD 344 344  
 CC FT CARBOHYD 365 365

FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 23 23 T -> S.  
 FT VARIANT 28 28 N -> A.  
 FT VARIANT 61 61 S -> L.  
 FT VARIANT 70 70 K -> R.  
 FT VARIANT 120 120 Y -> H.  
 FT VARIANT 125 125 E -> Q.  
 FT VARIANT 127 127 E -> G.  
 FT VARIANT 129 129 R -> H.  
 FT VARIANT 132 132 M -> V.  
 FT VARIANT 145 145 S -> N.  
 FT VARIANT 153 153 L -> V.  
 FT VARIANT 203 203 N -> D.  
 FT VARIANT 261 261 G -> R.  
 FT VARIANT 302 302 F -> L.  
 FT VARIANT 316 316 F -> L.  
 FT VARIANT 337 337 G -> R.  
 FT VARIANT 367 367 S -> N.  
 FT VARIANT 422 422 Q -> L.  
 FT VARIANT 429 429 I -> F.  
 FT VARIANT 433 433 D -> E.  
 SQ SEQUENCE 506 AA; 55093 MM; 6878310677FC9CB CRC64;  
 Query Match 7.0%; Score 188.5; DB 1; Length 506;  
 Best Local Similarity 22.4%; Pred. No. 6.3e-05;  
 Matches 71; Conservative 52; Mismatches 139; Indels 55; Gaps 11;  
 Oy 158 PRGNITGGK-----TLVSQLELDSGTWTCTVQNOQK--VERKIDVPPAPAPK 208  
 Db 88 PRVTNSDATKRNKMDISIRSNITTPADAGYVCVKRKERGMERK-----SGPETH 141  
 Oy 209 SCDKTHCPPELLGGPSVFLFPPPKDTLMISRPPEVTCVVVDVSH--EDPEVKFNMYVDG 266  
 Db 142 LTVAKSKSPFLSLSPV-----RATPEQVNFCTCHGSPRAISLKWPNFG 188  
 Oy 267 VEVNNAKTPREQYNSTYRVSVLYLHODWLNGKEKCKVSKALPA--PIEKTSKA 324  
 Db 189 NELASQTSVDPEDNNVSYSINSTTKVLATGVHSGQVTCVAHTLQGGPPL----- 241  
 Oy 325 KGGREQVYTLPPSRBELT-----KNVSLTCLYKGFYSPDIAVNESNQPPENNYKT 379  
 Db 242 RGTNLSETRVPEPTL-ETIGSPGAGQVNVTCQVNFYPMHQLTLWLENGMSRTEAS 300  
 Oy 380 PVLVDSGFPLYSKLTVDKSRMOQGVFSCVWHEALHNHYTQKSLSLPGLODETCA 439  
 Db 301 VFVANKGTFRQTSWFLVNSASRAHEAVVLTQVVEHDC-----QPAVSKNHTLEVSAPQ 354  
 Oy 440 EAQDEL----DGLWTT 452  
 Db 355 DQDTGTGTFPNDNWSWTS 371  
 RESULT 63  
 LAC1 MOUSE STANDARD; PRT; 105 AA.  
 ID LAC1 MOUSE  
 AC P01843;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ig lambda d-1 chain C region.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=83014953; PubMed=6812053;  
 RX Seising B., Miller J., Wilson R., Scorb U.,  
 RT "Evolution of mouse immunoglobulin lambda genes.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:4681-4685(1982).  
 RN [2]  
 RP SEQUENCE FROM N.A. (MOPC 315).  
 RX MEDLINE=81148806; PubMed=6259534;

RA Bothwell A.L.M., Paskind M., Schwartz R.C., Sonenshein G.E.,  
 RA Gelfer M.L., Baltimore D.;  
 RT "Dual expression of lambda genes in the MOPC-315 plasmacytoma";  
 RL Nature 290:65-67(1981).  
 RN (3)  
 RN SEQUENCE FROM N.A. (S43).  
 RX MEDLINE=62220143; PubMed=6283385;  
 RA Bothwell A.L.M., Paskind M., Rech M., Imanishi-Kari T., Rajewsky K.,  
 RA Baltimore D.;  
 RT "Somatic variants of murine immunoglobulin lambda light chains";  
 RL Nature 298:380-382(1982).  
 RN (4)  
 RN SEQUENCE (MYELOMA PROTEIN MOPC 104E).  
 RX MEDLINE=71107854; PubMed=5276767;  
 RA Appella E.;  
 RT "Amino acid sequences of two mouse immunoglobulin lambda chains";  
 RL Proc. Natl. Acad. Sci. U.S.A. 68:590-594(1971).  
 CC -1- MISCELLANEOUS: The MOPC 315 cell line produces 2 light chains, 1  
 CC normal lambda-2 chain and 1 abnormal lambda-1 chain that is  
 CC missing a large part of the V region. The C region sequence (shown  
 CC here) appears completely normal.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: J00582; AAA51636.1; -;  
 CC EMBL: J00587; AAB59672.1; -;  
 CC PIR: A93922; LIMS.  
 CC PDB: 1JNH; 06-FEB-02.  
 CC InterPro: IPR007110; IG-like.  
 CC InterPro: IPR003597; IG CL.  
 CC InterPro: IPR003006; IG\_MHC.  
 CC Pfam: PRF0047; Ig\_1.  
 CC SMART: SM00407; IgC1; 1.  
 CC PROSITE: PS50835; IG\_LIKE; 1.  
 CC PROSITE: PS00290; IG\_MHC; 1.  
 CC KM Immunoglobulin domain; Immunoglobulin C region; 3D-structure.  
 CC FT NON\_TER 1 1  
 CC FT DOMAIN 6 100 IG-LIKE.  
 CC FT DISULFID 27 86  
 CC FT DISULFID 104 104 INTERCHAIN (WITH HEAVY CHAIN).  
 CC FT CONFLICT 19 20 ET -> TE (IN REF. 4).  
 CC FT CONFLICT 56 56 O -> E (IN REF. 4).  
 CC FT CONFLICT 75 75 MISSING (IN REF. 4).  
 CC FT CONFLICT 81 82 HS -> SH (IN REF. 4).  
 CC FT CONFLICT 85 85 S -> SS (IN REF. 4).  
 CC FT CONFLICT 96 96 E -> Q (IN REF. 4).  
 CC FT SEQUENCE 105 AA; 11575 MW; A89F2B09BCFCA018 CRC64;  
 SO  
 Query Match 6.6%; Score 177.5; DB 1; Length 105;  
 Beest Local Similarity 35.9%; Pred. No. 4,4e-05;  
 Matches 37; Conservative 18; Mismatches 43; Indels 5; Gaps 3;  
 Oy 327 QPR-EPQVYTLPPSRDELTKGVSLTCLVKGFFPSDIAVWESNGQP-ENNYKTPPVLD 384  
 Db 1 QPKSSPSVTLFPSSSELELNKATLVCTITDFPGVTVDMKDGTPVYMGMTTPDSKO 60  
 Oy 385 SDGSFPLYSLKLVQDKSRMGQGNVFSGVMEALHNHYTKSL 427  
 Db 61 SNKYMAYSLYTLTARABRHSYSQVTHE--GHTVERSL 100  
 RESULT 64  
 VCAL\_RAT  
 ID VCAL\_RAT STANDARD; PRT; 739 AA.  
 AC P29534;  
 DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Vascular cell adhesion protein 1 precursor (V-CAM 1).  
 GN VCAM1 OR VCAM-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN (1)  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Lung;  
 RC MEDLINE=92181437; PubMed=1371918;  
 RX Hession C., Moy P., Tizard R., Chisholm P., Williams C., Wysk M.,  
 RA Burky L., Miyake K., Kincaid P., Lobb R.;  
 RT "Cloning of murine and rat vascular cell adhesion molecule-1";  
 RL Biochem. Biophys. Res. Commun. 183:163-169(1992).  
 CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION  
 CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1  
 CC INTEGRIN VIA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL  
 CC TRANSDUCTION. THE VCAM1/VIA4 INTERACTION MAY PLAY A  
 CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE  
 CC EMIGRATION TO SITES OF INFLAMMATION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as  
 CC well as on macrophage-like and dendritic cell types in both normal  
 CC and inflamed tissue.  
 CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M8488; AAA42332.1; -;  
 CC PIR: J50675; J50675.  
 CC HSSP: P19320; 1VCA.  
 CC InterPro: IPR003987; ICAM\_VCAM-1.  
 CC InterPro: IPR007110; IG-like.  
 CC InterPro: IPR003598; IG C2.  
 CC InterPro: IPR003989; VCAM-1.  
 CC Pfam: PRF0047; Ig\_5.  
 CC PRINTS: PR01472; ICAMVCAM1.  
 CC DR PRINTS; PRO1474; VCAM1.  
 CC SMART: SM00408; IG\_C2; 3.  
 CC PROSITE: PS50835; IG\_LIKE; 5.  
 CC KM Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;  
 CC Repeat; Signal.  
 CC FT SIGNAL 1 24  
 CC FT CHAIN 25 739  
 CC FT DOMAIN 25 698  
 CC FT TRANSMEM 699 720  
 CC FT DOMAIN 721 739  
 CC FT DOMAIN 25 111  
 CC FT DOMAIN 119 212  
 CC FT DOMAIN 223 309  
 CC FT DOMAIN 312 397  
 CC FT DOMAIN 408 506  
 CC FT DOMAIN 514 595  
 CC FT DOMAIN 601 682  
 CC FT DISULFID 47 95  
 CC FT DISULFID 52 99  
 CC FT DISULFID 137 195  
 CC FT CARBOHYD 273 273  
 CC FT CARBOHYD 424 424  
 CC FT CARBOHYD 531 531  
 CC FT CARBOHYD 561 561  
 CC FT CARBOHYD 650 650  
 CC FT SEQUENCE 739 AA; 81246 MW; 5C6085A1A1B100C CRC64;  
 SO  
 Query Match 6.5%; Score 176.5; DB 1; Length 739;  
 FT SIGNAL 1 24  
 FT CHAIN 25 739  
 FT DOMAIN 25 698  
 FT TRANSMEM 699 720  
 FT DOMAIN 721 739  
 FT DOMAIN 25 111  
 FT DOMAIN 119 212  
 FT DOMAIN 223 309  
 FT DOMAIN 312 397  
 FT DOMAIN 408 506  
 FT DOMAIN 514 595  
 FT DOMAIN 601 682  
 FT DISULFID 47 95  
 FT DISULFID 52 99  
 FT DISULFID 137 195  
 FT CARBOHYD 273 273  
 FT CARBOHYD 424 424  
 FT CARBOHYD 531 531  
 FT CARBOHYD 561 561  
 FT CARBOHYD 650 650  
 FT SEQUENCE 739 AA; 81246 MW; 5C6085A1A1B100C CRC64;  
 SO  
 Query Match 6.5%; Score 176.5; DB 1; Length 739;  
 FT SIGNAL 1 24  
 FT CHAIN 25 739  
 FT DOMAIN 25 698  
 FT TRANSMEM 699 720  
 FT DOMAIN 721 739  
 FT DOMAIN 25 111  
 FT DOMAIN 119 212  
 FT DOMAIN 223 309  
 FT DOMAIN 312 397  
 FT DOMAIN 408 506  
 FT DOMAIN 514 595  
 FT DOMAIN 601 682  
 FT DISULFID 47 95  
 FT DISULFID 52 99  
 FT DISULFID 137 195  
 FT CARBOHYD 273 273  
 FT CARBOHYD 424 424  
 FT CARBOHYD 531 531  
 FT CARBOHYD 561 561  
 FT CARBOHYD 650 650  
 FT SEQUENCE 739 AA; 81246 MW; 5C6085A1A1B100C CRC64;  
 SO

Best Local Similarity 21.4%; Pred. No. 0.00059;  
Matches 101; Conservative 71; Mismatches 169; Indels 131; Gaps 23;

```

QY 33 KGDYELTQTASQKSIQPHMKNQSIKILNGQSEFLTGPBKLNDRADRSRLMDQNF 92
DB 238 EAAAVTMTCASGLPAPELFWSK-----LDNGVQLL-----SSNA 274
QY 93 PLIKNLKIEDSDTYICE-----VEDQKEVQLLV-----FGLTANSPDTHLLQGLT 142
DB 275 TLTLLAMREDSDGYICEGVNLVGDKTEVLEIVQEKPTVDISGQVAAQGVSVLT 334
QY 143 LESPGSSPSVQCRSPRGKNIG-----GKTLVSQLELDSDGTWTCTVLQNKVYF 195
DB 335 CAACVCDSPSFGMRQTDSPLNGEVRDEGATSTLTLSIPGVDEHSHYLTCTVQGRKLEK 394
QY 196 KLDIYPCAPAPKPSCKDKHTHCEPL-LGGSVFLFPKPKDPTLMTIRTEYVCVVYD--- 251
DB 395 TLQVEYVSFPD-----PEIEISGP-----LVHGRPTVNTCTVENYVPF 433
QY 252 SHEDEEV-----KFNWYVDGVEVHNAKTKPREEOY-----NSTYRVSVLTVLHODW 298
DB 434 DLELELKLKGETTLNKF--LAEIGTSLFETKSLFEMPIFAEDTGALVCLATHSSQ 491
QY 299 LNKGYKCKVSNKALPAPIEKTISRAKQGPPEQVYLP-PSRDELTKNQVSLTGLVKG 357
DB 492 ME-SEPKQKQSTQTLVYVNV-----APKEPTIWSPSPVPER--GSPVNLTCSSDGF 539
QY 358 YPSDAVEME---SNGQPENNYKTPPVLDSDGSFVLSKLTVDKSRNQGVFSCVTH 414
DB 540 PTPKLL--WSRLKNKEQLQ-----PLSQNTTSLFPMATKM-----EDSGIYVCEGIN 583
QY 415 EALAHNYTKSLSL-----SPGLQF-----DETCAEAGDGLGLM 450
DB 584 EA---GISKKSVELLIQSSKDIQLTVFPPSKSVKGDVYIISCTGSGVPEIM 632

```

```

RESULT 65
ID PGHM HUMAN STANDARD; PRT; 4391 AA.
AC P98160; O16287; Q9H3V5;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (perlecan) (PLC).
GN HSPG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RX MEDLINE=92112994; PubMed=1730768;
RA Kallunki P., Tryggvason K.;
RT "Human basement membrane heparan sulfate proteoglycan core protein: a
RT 467-kD protein containing multiple domains resembling elements of the
RT low density lipoprotein receptor, laminin, neural cell adhesion
RT molecules, and epidermal growth factor."
RL J. Biol. Chem. 267:8544-8557(1992).
RN 12
RX MEDLINE=92235084; PubMed=1569102;
RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT "Primary structure of the human heparan sulfate proteoglycan from
RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT domains homologous to the low density lipoprotein receptor, laminin,
RT neural cell adhesion molecules, and epidermal growth factor."
RL J. Biol. Chem. 267:8544-8557(1992).
RN 13
RX MEDLINE=20553141; PubMed=1101850;
RA Nicole S., Davoine C.-S., Topaloglu H., Catolico L., Barral D.,

```

```

RA Beighton P., Ben-Hamida C., Hammouda H., Craud C., White P.S.,
RA Sanson D., Urtizberea J.A., Lehmann-Horn F., Weisenbach J.,
RA Hentati F., Fontaine B., Lehmann-Horn F., Weisenbach J.,
RT "Perlecan, the major proteoglycan of basement membranes, is altered in
RT patients with Schwartz-Jampel syndrome (chondrodysdystrophic myotonia).";
RL Nat. Genet. 26:480-483(2000).
RN 14
RX MEDLINE=1016-1470 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=91365376; PubMed=1679749;
RA Dodge G.R., Kovalszky I., Chu M.L., Hassell J.R., McBride O.W.,
RA Yi H.F., Iozzo R.V.;
RT "Heparan sulfate proteoglycan of human colon: partial molecular
RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
RT short arm of human chromosome 1.";
RL Genomics 10:673-680(1991).
RN 15
RX MEDLINE=890-1396 FROM N.A.
RC TISSUE=Fibrocarcoma;
RX MEDLINE=92120660; PubMed=1685141;
RA Kallunki P., Eddy R.L., Byers M.G., Keestila M., Shove T.B.,
RA Tryggvason K.;
RT "Cloning of human heparan sulfate proteoglycan core protein,
RT assignment of the gene (HSPG2) to 1p36.1-1p35 and identification of
RT a BamHI restriction fragment length polymorphism.";
RL Genomics 11:389-396(1991).
RN 16
RX MEDLINE=94052171; PubMed=8234307;
RA Cohen I.R., Graessell S., Murdoch A.D., Iozzo R.V.;
RT "Structural characterization of the complete human perlecan gene and
RT its promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
RN 17
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666(2003).
RN 18
RX MEDLINE=92112994; PubMed=1730768;
RA Kallunki P., Tryggvason K.;
RT "Human basement membrane heparan sulfate proteoglycan core protein: a
RT 467-kD protein containing multiple domains resembling elements of the
RT low density lipoprotein receptor, laminin, neural cell adhesion
RT molecules, and epidermal growth factor."
RL J. Biol. Chem. 267:8544-8557(1992).
RN 19
RX MEDLINE=92235084; PubMed=1569102;
RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT "Primary structure of the human heparan sulfate proteoglycan from
RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT domains homologous to the low density lipoprotein receptor, laminin,
RT neural cell adhesion molecules, and epidermal growth factor."
RL J. Biol. Chem. 267:8544-8557(1992).
RN 20
RX MEDLINE=20553141; PubMed=1101850;
RA Nicole S., Davoine C.-S., Topaloglu H., Catolico L., Barral D.,

```

```

DR EMBL; X62515; CAA44373.1; -.
DR EMBL; M85289; AAA52700.1; -.
DR EMBL; AL445795; CAC18534.1; -.
DR EMBL; M64283; AAA52699.1; -.
DR EMBL; S76436; AAB2121.2; -.
DR EMBL; L22078; -, NOT_ANNOTATED_CDS.
DR PIR; A38096; A38096.
DR HSP; P00740; IEDM.
DR Siema-2DPAGE; P98160; -.
DR GeneW; HGNC:3273; HSP62.
DR MIM; 142461; -.
DR MIM; 255800; -.
DR InterPro; IPR008985; ConA_like_1ec_g1.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_1like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003596; Ig_V.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PR000082; SEA_domain.
DR Pfam; PR00008; EGF_4.
DR Pfam; PR00047; Ig_22.
DR Pfam; PR00052; Laminin_B_3.
DR Pfam; PR00053; Laminin_EGF_7.
DR Pfam; PR00054; Laminin_G_3.
DR Pfam; PR00057; Idl_recept_e_4.
DR Pfam; PR01390; SEA_1.
DR PRINTS; PR00261; LDLRECEPTOR.
DR ProDom; PD001031; Laminin_B_3.
DR SMART; SM00181; EGF_15.
DR SMART; SM00180; EGF_Lam; 12.
DR SMART; SM00409; Ig_22.
DR SMART; SM00408; IGC2; 21.
DR SMART; SM00406; IGV_7.
DR SMART; SM00281; Lamb; 3.
DR SMART; SM00282; Lamg; 3.
DR SMART; SM00192; LDLa; 4.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS50026; EGF_3; 4.
DR PROSITE; PS50835; IG_Like; 22.
DR PROSITE; PS50025; LAM_G_DOMAIN; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF_11.
DR PROSITE; PS01209; LDLRA_1; 4.
DR PROSITE; PS50068; LDLRA_2; 4.
DR PROSITE; PS50024; SEA; 1.
DR Signal; Baement membrane; Proteoglycan; Repeat; Glycoprotein;
KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
KW Extracellular matrix; EGF-like domain; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 4391
FT FT 80 194
FT DOMAIN 198 235
FT DOMAIN 284 320
FT DOMAIN 324 360
FT DOMAIN 367 404
FT DOMAIN 405 504
FT DOMAIN 521 530
FT DOMAIN 531 730
FT DOMAIN 731 763
FT DOMAIN 764 813
FT DOMAIN 814 871
FT DOMAIN 879 923
FT DOMAIN 924 933
FT DOMAIN 934 1125
FT DOMAIN 1126 1156

```

```

FT DOMAIN 1159 1208 LAMININ EGF-LIKE 6.
FT DOMAIN 1209 1265 LAMININ EGF-LIKE 7.
FT DOMAIN 1275 1324 LAMININ EGF-LIKE 8.
FT DOMAIN 1325 1334 LAMININ EGF-LIKE 9 (N-TERMINAL).
FT DOMAIN 1335 1529 LAMININ DOMAIN IV 3 (DOMAIN III C).
FT DOMAIN 1530 1562 LAMININ EGF-LIKE 9 (C-TERMINAL).
FT DOMAIN 1563 1612 LAMININ EGF-LIKE 10.
FT DOMAIN 1613 1670 LAMININ EGF-LIKE 11.
FT DOMAIN 1671 1771 LAMININ EGF-LIKE 12.
FT DOMAIN 1772 1865 LAMININ EGF-LIKE 13.
FT DOMAIN 1866 1955 LAMININ EGF-LIKE 14.
FT DOMAIN 1956 2051 LAMININ EGF-LIKE 15.
FT DOMAIN 2052 2151 LAMININ EGF-LIKE 16.
FT DOMAIN 2152 2244 LAMININ EGF-LIKE 17.
FT DOMAIN 2245 2340 LAMININ EGF-LIKE 18.
FT DOMAIN 2341 2436 LAMININ EGF-LIKE 19.
FT DOMAIN 2437 2533 LAMININ EGF-LIKE 20.
FT DOMAIN 2534 2629 LAMININ EGF-LIKE 21.
FT DOMAIN 2630 2726 LAMININ EGF-LIKE 22.

Query Match 6.5%; Score 175.5; DB 1; Length 4391;
Best Local Similarity 20.7%; Pred. No. 0.0064;
Matches 100; Conservative 67; Mismatches 194; Indels 123; Gaps 22;

QY 33 KSDTVELCTASQKSIQPHMKNNSQIKILNGSGFLFKGPKLMDRADSRSLMDQGNF 92
Db 2740 ESETLDLNCVVPQMAQVTH-----KRGSL--PSHHOTRGRRL----- 2779
QY 93 PLIRKLKIEDSDTYICEVEDK--EEVOLVFGILTANSDTH----- 132
Db 2780 ---LHVSPADSGEYVVCVWSSGSLASVLTITASGSAHVPAFGAPRIEPPSS 2836
QY 133 -LLQGSLTLTIESPPSSPPVQCSPPKNIQG-----GTLVSQLELDGCTWTCTV 186
Db 2837 RVAEQTLIDLC-VVPGQAHQVTHKRGKGLPARQVHGPLRLNQVSPADSGEYSCV 2895
QY 187 LONOKVPEKIDV-----PCPAPPKSCDKHTKTCPELLGSGSVLPFPKPKDTLMISRT 241
Db 2896 TGSSGTLASVLTIEPPSPGPIPA-----GLAQTYIEASSHYTEGT 2941
QY 242 FEVTCVVVDVSHEDPEVFNFVVDGEVHNATKPREQYNSTYRVSVLTVLHODWING 301
Db 2942 LDINCVVPGQH---AQVTWYKRG-----GSLPRHQTHSQRLHLVSPA-----DS 2986
QY 302 KEYCKVANKALP--APIEKTISKAKQP--REPQVYTLPPSDLTAKNOVSLTCLV- 354
Db 2987 GEYVCRASGPPEQGEAFVTVPPESSGYRLRSPVISIDPPSSSTVOQGDASFCLIH 3046
QY 355 KGFVPSDIAVESNSGQF--ENNYKTPPVLDSDGFFLYSKLTVDKSRMQGNVPSG--- 410
Db 3047 DGAAP--ISLEWTKNQLIEDVHISP-----NGSI-----ITVGTSPSNHGTVCVAS 3094
QY 411 -----SYMHEALNHYTKSLISLP-----GLQDTECAEADGELDGLTTDPPRAS 458
Db 3095 NAYGAGQSVNLSVGPFTVSVLPFGPVVKKYKAVTIECVSAGSPRSARAT-----HIS 3150
QY 459 ALPA 462
Db 3151 STPA 3154

RESULT 66
LAC2_RAT STANDARD; PRT; 104 AA.
AC P20767;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda-2 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OK NCBI_Taxid=10116;

```





CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X59350; CAA42006.1; -  
 DR EMBL; U62631; AAB06448.1; -  
 DR EMBL; U62631; AAB06449.1; -  
 DR EMBL; X52785; CAA36988.1; ALT\_FRAME.  
 DR EMBL; AB012996; BAA36565.1; -  
 DR EMBL; AB012997; BAA36566.1; -  
 DR EMBL; AB012998; BAA36567.1; -  
 DR EMBL; AB012999; BAA36568.1; -  
 DR EMBL; AB013000; BAA36569.1; -  
 DR EMBL; AB013002; BAA36571.1; -  
 DR EMBL; AB013003; BAA36572.1; -  
 DR EMBL; AB013004; BAA36573.1; -  
 DR EMBL; AB013006; BAA36575.1; -  
 DR Gene; HGNC:1643; CD22.  
 DR MIM; 107266; -  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR Pfam; PF00047; Ig\_6.  
 DR SMART; SMO0408; IGC2\_4.  
 DR PROSITE; PS00835; IG\_LIKE; 6.  
 KW Cell adhesion; lectin; Antigen; Transmembrane; Signal; Glycoprotein;  
 KW Immunoglobulin domain; Repeat; Phosphorylation; Alternative splicing;  
 KW Polymorphism.  
 FT SIGNAL 1 19  
 FT CHAIN 20 847  
 FT DOMAIN 20 687  
 FT TRANSMEM 688 706  
 FT DOMAIN 707 847  
 FT DOMAIN 20 138  
 FT DOMAIN 143 235  
 FT DOMAIN 242 326  
 FT DOMAIN 331 416  
 FT DOMAIN 419 500  
 FT DOMAIN 505 582  
 FT DOMAIN 593 676  
 FT SITE 760 795  
 FT SITE 794 799  
 FT SITE 820 825  
 FT SITE 840 845  
 FT DISULFID 39 167  
 FT DISULFID 44 102  
 FT DISULFID 161 219  
 FT DISULFID 265 309  
 FT DISULFID 353 396  
 FT DISULFID 442 484  
 FT DISULFID 529 571  
 FT DISULFID 616 659  
 FT MOD\_RES 762 762  
 FT MOD\_RES 807 807  
 FT MOD\_RES 822 822  
 FT MOD\_RES 842 842  
 FT CARBOHYD 67 67  
 Query Match 6.2%; Score 167; DB 1; Length 847;  
 Best Local Similarity 21.2%; Pred. No. 0.0028;  
 Matches 95; Conservative 70; Mismatches 154; Indels 130; Gaps 23;  
 32 KKGDTVELCTASQKKSIOFHKNKSNQIKILGNQGSFLTKGPSKLNDRADRSRLMDQGN 91

Db :::::|:|:|-----SNPEYTV-----SWLKDG-----TSKQNT 291  
 256 REGSYMTCEVSS-----  
 Qy 92 FPLIKLKLIKEDSTPTICEVED-----QKEEVOLLYVGLTANSDTHL-----CQOS 138  
 Db 292 FTLLNLREVTYKXQSKYCCQVSNVDPGRSEVFLQVYAPRPSTVQLISPAVSGOYEF 351  
 Qy 139 LTLTLEPPGSPSPVQCRSPRGKNIQ--GKTLVSLSLELDSDSGTWCTVQLONQKVEFK 196  
 Db 352 LQMSLAP---LPNTNYWYNGKEMQRTBEKHIPIILPHNAHTYGC-VAEN-----400  
 Qy 197 IDIVPCAPPEPKSCDKHTHCPPELLG-----GPSVFL---PPPKPDTLMISFTP---EVT 245  
 Db 401 -----ILGTQGRPGALDQVPPKATTVIQLNMPPIREGDTV 438  
 Qy 246 CVVVDVSHEDPEV-KFWVYVDGVFNHAKTKPREQNSYTRVSVLTVLHQDLNGKEY 304  
 Db 439 TILSCVYNSNPSVTRYEW-----KPGHAWEPSS-----LGVLTQVWGDN-TTI 482  
 Qy 305 KCKVSNK-ALPAPIEKTIKAKQPREPQVYTLPPRDELTKNQVSLTCLVKGFPSPDI 362  
 Db 483 ACAACNWCMSWASVVALNQYA---PRDVYRKIKPLSEIHSGSVSLQCDPSSSHREV 539  
 Qy 363 AVENESNG-----QENNYKTPPEVLDSGFFLYSLTYDKSRMQGANVFCSVMEHAL 417  
 Db 540 QFPEKNGRLLGKESQLNFDPSISP--EDAGSY-----SCVANNISIGQTASKAWTL 587  
 Qy 418 HNHYTQSL--SLSPGIQDDE-----TC 438  
 Db 588 EVLYAPRLRVSMSPGQVMEKGSATLTC 616

RESULT 69  
 SRBL\_HUMAN STANDARD; PRT; 398 AA.  
 ID SRBL\_HUMAN 000241; Q8TB12; Q9H1U5; Q9Y4V0;  
 AC 28-FEB-2003 (Rel. 41, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Signal-regulatory protein beta-1 precursor (SRP-beta-1).  
 GN SRPB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RX MEDLINE=97215901; PubMed=9062191;  
 RA Kharitonov A., Chen Z., Sures I., Wang H., Schilling J.,  
 RA Ullrich A.;  
 RT "A family of proteins that inhibit signalling through tyrosine kinase  
 RT receptors.";  
 RL Nature 386:181-186(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagunley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Buttrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Leiva-Salido M.H., Leverkus M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,



RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showken R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,  
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.,  
 RT "The DNA sequence and comparative analysis of human chromosome 20.",  
 RL Nature 414:865-871(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Skin;  
 RA MEDLINE=22386257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heist F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Cairnci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield J.S.N., Krzywinski M.I., Skalek U., Smalleg D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP INTERACTION WITH TYROBP AND SYK.  
 RA PubMed=11169422;  
 RA Tomaseello E., Cant C., Buehring H.-J., Vely F., Andre P., Seifert M.,  
 RA Ullrich A., Vavler E.,  
 RT "Association of signal-regulatory proteins beta with KARAP/DAP-12.",  
 RL Eur. J. Immunol. 30:2147-2156(2000).  
 RN [5]  
 RP INTERACTION WITH TYROBP, N-GLYCOSYLATION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=20072721; PubMed=10604985;  
 RA Dietrich J., Cella M., Seifert M., Buehring H.-J., Colonna M.,  
 RT "Signal-regulatory protein beta 1 is a DAP12-associated activating  
 RT receptor expressed in myeloid cells.",  
 RL J. Immunol. 164:9-12(2000).  
 CC -1- FUNCTION: Immunoglobulin-like cell surface receptor involved in  
 CC the negative regulation of receptor tyrosine kinase-coupled  
 CC signaling processes. Participates also in the recruitment of  
 CC tyrosine kinase SYK.  
 CC -1- SUBUNIT: Interacts with TYROBP. This interaction results in the  
 CC recruitment of SYK.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=O00241-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=O00241-2; Sequence=VSP\_007026;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Detected in monocytes and dendritic cells.  
 CC -1- PTM: N-glycosylated.  
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Y10376; CAA71404.1; -  
 CC EMBL: AL049634; CAB4661.2; -  
 CC EMBL: AL13804; CAC17540.1; -  
 CC EMBL: BC025286; AAH25286.1; -  
 CC Genew: HGNC:15928; STRP1.  
 CC MIM: 603889; -  
 CC GO: GO:0005887; C: integral to plasma membrane; TAS.  
 CC GO: GO:0007166; P: cell surface receptor linked signal transdu. .; TAS.  
 CC InterPro: IPR007110; IG-1-like.  
 CC InterPro: IPR003597; IG\_C1.  
 CC Pfam: PF00047; Ig\_3.  
 CC SMART: SM00407; Ig\_3.  
 CC PROSITE: PS00835; IG\_LIKE; 3.  
 CC Repeat; Signal; Transmembrane; Immunoglobulin domain; Glycoprotein;  
 CC Alternative splicing.  
 CC FT SIGNAL 1 26  
 CC FT CHAIN 27 398  
 CC FT DOMAIN 27 371  
 CC FT TRANSMEM 372 392  
 CC FT DOMAIN 393 398  
 CC FT DOMAIN 27 136  
 CC FT DOMAIN 147 246  
 CC FT DOMAIN 253 347  
 CC FT DISULFID 34 120  
 CC FT DISULFID 169 227  
 CC FT DISULFID 272 330  
 CC FT DISULFID 244 244  
 CC FT CARBOHYD 244 244  
 CC FT CARBOHYD 269 269  
 CC FT CARBOHYD 291 291  
 CC FT VARSPLIC 145 361  
 CC FT CONFLICT 53 53  
 CC FT CONFLICT 102 102  
 CC FT CONFLICT 229 229  
 CC FT CONFLICT 363 363  
 CC SQ SEQUENCE 398 AA; 43255 NM; A2AA08FE8B2BC52B CRC64;  
 CC  
 CC Query Match 6.2%; Score 166.5; DB 1; Length 398;  
 CC Best Local Similarity 24.2%; Pred. No. 0.0012;  
 CC Matches 82; Conservative 50; Mismatches 154; Indels 53; Gaps 14;  
 CC  
 CC QY 136 GGSLLTLESPPGSSBSPVQCSPPKRNIGGKTLVSQLELDSGTWTTLVQ--NQKV 193  
 CC DB 72 GAGRELIYNQKKGHPRTVTVSELTRNNLDFISISINITPADAGTYCVCKFRKSPDDV 131  
 CC QY 194 EKKIDVPCPAPPEPSCDKTHCPRLGSPSYFLPPPKKDTLMSRPEVYCVVVDVSH 253  
 CC DB 132 EFK----SGAGTLESVRKPPAP--VSGPAP-----RATPEHTVSTCESH 172  
 CC QY 254 --EDPEKFNWVVDGVENVNAKTK--PRE-----QYNSTRVVSVLVLDHDLNGKEYK 305  
 CC DB 173 GSPSPDITLKWPNKNGELSDPOTNVDPADGSYSIHSIRARVLRGDVHSGVI----- 226  
 CC QY 306 CYSNKAFLPA-DIEKT--ISKAGQPREQVYTLPPSRDELTKNVSLTCLVKGFPYSDI 362  
 CC DB 227 CEMAHITLQGDPLRGTANLSEAIRVPLEV--TQPMRAE---NNAVTCQVSNFPRGL 282  
 CC QY 363 AVWESNCPENNNYKTPPVLDSDGSFLYSLTLYDKSKWQGNPFGSCVWHEALHNHT 422  
 CC DB 283 QLTWLENGVSVSTETASTLIENKDDTYMMWSGLVNTCAHRDQVVLTCQVEHDG----- 336  
 CC QY 423 QKSLSPGLQDETCAEKNODELGLMTTPPRASALP 461  
 CC DB 337 QQAVSKYALRI-----SAHQEHGSDITHEPALPTAP 370  
 CC  
 CC RESULT 70  
 CC UN89 CAEEL STANDARD; PRT; 6632 AA.  
 CC ID UN89 CAEEL  
 CC AC 001761; Q17362;

DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).  
 GN UNC-89 OR C09D1.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.  
 RC STRAIN=Br18c01 N2;  
 RC MEDLINE=96180278; PubMed=8603916;  
 RA Benjan G.M., Tinley T.L., Tang X., Borodovsky M.;  
 RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line  
 RT assembly, encodes a giant modular protein composed of Ig and signal  
 RT transduction domains.";  
 RT J. Cell Biol. 132:835-848 (1996).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Br18c01 N2;  
 RC Du Z., Le T.T., Wilson R.;  
 RL Submitted (May-1997) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.  
 CC -! FUNCTION: Structural component of the muscle M-line. Myofibrilment  
 CC lattice assembly begins with positional cues laid down in the  
 CC basement membrane and muscle cell membrane. UNC-89 responds to  
 CC these signals, localizes, and then participates in assembling an  
 CC M-line.  
 CC -! TISSUE SPECIFICITY: Localizes to the middle of A-bands.  
 CC -! SIMILARITY: Contains 1 DBL-homology (DH) domain.  
 CC -! SIMILARITY: Contains 1 fibronectin type III domain.  
 CC -! SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.  
 CC -! SIMILARITY: Contains 1 PH domain.  
 CC -! SIMILARITY: Contains 5 RCD domain.  
 CC -! SIMILARITY: Contains 1 SH3 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U33058; AAB00542.1; -.  
 DR EMBL: AF003131; AAB54132.2; -.  
 DR PDB: 1PHO; 20-DEC-00.  
 DR WormPep: C09D1.1; CE30426.  
 DR InterPro: IPR008957; FN\_III-like.  
 DR InterPro: IPR003961; FN\_III.  
 DR InterPro: IPR007110; IG-I-like.  
 DR InterPro: IPR003598; IG\_C2.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR InterPro: IPR001849; PH.  
 DR InterPro: IPR007850; RCD.  
 DR InterPro: IPR000219; RHOGEF.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF000447; f3; 1.  
 DR Pfam: PF00047; f3; 1.  
 DR Pfam: PF00169; PH; 1.  
 DR Pfam: PF01777; RCD; 5.  
 DR Pfam: PF00621; RHOGEF; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SMO0408; IG\_C2; 23.  
 DR SMART: SMO0325; RHOGEF; 1.  
 DR SMART: SMO0326; SH3; 1.  
 DR PROSITE: PS50010; DH\_2; 1.  
 DR PROSITE: PS50835; IG-LIKE; 49.  
 DR PROSITE: PS50003; PH\_DOMAIN; 1.

DR PROSITE; PS50002; SH3; 1.  
 KW Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;  
 KW 3D-structure.  
 FT DOMAIN 63 127 SH3.  
 FT DOMAIN 152 330 DH.  
 FT DOMAIN 342 498 PH.  
 FT DOMAIN 547 633 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 648 736 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 748 838 IG-LIKE C2-TYPE 3.  
 FT DOMAIN 946 1033 IG-LIKE C2-TYPE 4.  
 FT DOMAIN 1044 1132 IG-LIKE C2-TYPE 5.  
 FT DOMAIN 1140 1227 IG-LIKE C2-TYPE 6.  
 FT DOMAIN 1272 1315 THR-RICH.  
 FT DOMAIN 1375 1475 RCD 1.  
 FT DOMAIN 1479 1585 RCD 2.  
 FT DOMAIN 1597 1695 RCD 3.  
 FT DOMAIN 1700 1799 RCD 4.  
 FT DOMAIN 1800 1860 RCD 5.  
 FT DOMAIN 1982 2067 IG-LIKE C2-TYPE 7.  
 FT DOMAIN 2071 2163 IG-LIKE C2-TYPE 8.  
 FT DOMAIN 2171 2261 IG-LIKE C2-TYPE 9.  
 FT DOMAIN 2269 2359 IG-LIKE C2-TYPE 10.  
 FT DOMAIN 2367 2455 IG-LIKE C2-TYPE 11.  
 FT DOMAIN 2463 2564 IG-LIKE C2-TYPE 12.  
 FT DOMAIN 2563 2651 IG-LIKE C2-TYPE 13.  
 FT DOMAIN 2657 2746 IG-LIKE C2-TYPE 14.  
 FT DOMAIN 2754 2858 IG-LIKE C2-TYPE 15.  
 FT DOMAIN 2887 2980 IG-LIKE C2-TYPE 16.  
 FT DOMAIN 2994 3081 IG-LIKE C2-TYPE 17.  
 FT DOMAIN 3087 3183 IG-LIKE C2-TYPE 18.  
 FT DOMAIN 3189 3280 IG-LIKE C2-TYPE 19.  
 FT DOMAIN 3286 3376 IG-LIKE C2-TYPE 20.  
 FT DOMAIN 3384 3469 IG-LIKE C2-TYPE 21.  
 FT DOMAIN 3482 3572 IG-LIKE C2-TYPE 22.  
 FT DOMAIN 3580 3667 IG-LIKE C2-TYPE 23.  
 FT DOMAIN 3686 3777 IG-LIKE C2-TYPE 24.  
 FT DOMAIN 3817 3908 IG-LIKE C2-TYPE 25.  
 FT DOMAIN 3920 4009 IG-LIKE C2-TYPE 26.  
 FT DOMAIN 4018 4106 IG-LIKE C2-TYPE 27.  
 FT DOMAIN 4109 4201 IG-LIKE C2-TYPE 28.  
 FT DOMAIN 4212 4297 IG-LIKE C2-TYPE 29.  
 FT DOMAIN 4302 4387 IG-LIKE C2-TYPE 30.  
 FT DOMAIN 4400 4485 IG-LIKE C2-TYPE 31.  
 FT DOMAIN 4489 4580 IG-LIKE C2-TYPE 32.  
 FT DOMAIN 4588 4678 IG-LIKE C2-TYPE 33.  
 FT DOMAIN 4681 4771 IG-LIKE C2-TYPE 34.  
 FT DOMAIN 4873 4961 IG-LIKE C2-TYPE 35.  
 FT DOMAIN 4965 5057 IG-LIKE C2-TYPE 36.  
 FT DOMAIN 5067 5160 IG-LIKE C2-TYPE 37.  
 FT DOMAIN 5171 5260 IG-LIKE C2-TYPE 38.  
 FT DOMAIN 5277 5366 IG-LIKE C2-TYPE 39.  
 FT DOMAIN 5383 5472 IG-LIKE C2-TYPE 40.  
 FT DOMAIN 5487 5578 IG-LIKE C2-TYPE 41.  
 FT DOMAIN 5595 5685 IG-LIKE C2-TYPE 42.  
 FT DOMAIN 5701 5790 IG-LIKE C2-TYPE 43.  
 FT DOMAIN 5815 5904 IG-LIKE C2-TYPE 44.  
 FT DOMAIN 5925 6014 IG-LIKE C2-TYPE 45.  
 FT DOMAIN 6038 6130 IG-LIKE C2-TYPE 46.  
 FT DOMAIN 6150 6239 IG-LIKE C2-TYPE 47.  
 FT DOMAIN 6275 6368 FIBRONECTIN TYPE-III.  
 FT DOMAIN 6413 6502 IG-LIKE C2-TYPE 48.  
 FT DOMAIN 6507 6596 IG-LIKE C2-TYPE 49.  
 FT DISULFID 568 621 POTENTIAL.  
 FT DISULFID 2908 2975 POTENTIAL.  
 FT DISULFID 3015 3065 POTENTIAL.  
 FT DISULFID 3707 3759 POTENTIAL.  
 FT DISULFID 3826 3890 POTENTIAL.  
 FT DISULFID 5092 5157 POTENTIAL.  
 FT DISULFID 5298 5350 POTENTIAL.  
 FT DISULFID 5508 5560 POTENTIAL.  
 FT DISULFID 5616 5669 POTENTIAL.  
 FT DISULFID 5722 5764 POTENTIAL.  
 FT DISULFID 5836 5901 POTENTIAL.

```

FT DISULFID 5946 5998 POTENTIAL.
FT DISULFID 6036 6171 POTENTIAL.
FT DISULFID 6421 6486 POTENTIAL.
FT CONFLICT 2137 2137 A -> P (IN REF. 1).
FT CONFLICT 2245 2247 A -> P (IN REF. 1).
FT CONFLICT 2258 2258 A -> P (IN REF. 1).
FT CONFLICT 2284 2284 M -> G (IN REF. 1).
FT CONFLICT 2297 2297 E -> I (IN REF. 1).
FT CONFLICT 3531 3531 A -> G (IN REF. 1).
FT CONFLICT 3884 3888 DAGEY -> RRRRI (IN REF. 1).
FT CONFLICT 3929 3929 A -> V (IN REF. 1).
FT CONFLICT 5134 5134 A -> P (IN REF. 1).
FT CONFLICT 5145 5145 T -> S (IN REF. 1).
FT CONFLICT 5185 5185 G -> A (IN REF. 1).
FT CONFLICT 5199 5199 K -> N (IN REF. 1).
FT CONFLICT 5202 5202 L -> F (IN REF. 1).
FT CONFLICT 5213 5213 F -> L (IN REF. 1).
FT CONFLICT 6178 6178 A -> G (IN REF. 1).
FT CONFLICT 6268 6268 K -> E (IN REF. 1).
SQ SEQUENCE 6632 AA; 73165 MW; 26203EDD62960B89 CRC64;

Query Match
Best Local Similarity 23.6%; Score 164; DB 1; Length 6632;
Matches 107; Conservative 59; Mismatches 175; Indels 112; Gaps 22;

QY 27 KVLVGGKGDVLTCTASQKKSIOF-HMNSNQIKLNGSPFLTGSPKLNDRADSRSS 85
DB 3693 KETPAVEGDTVLECKVKNKESHPOIKFKNQDPVEL---GQHM-----QKE 3735
QY 86 LMDQGNFPLIKNLKIEDSDTYICE-----VED-----QKEE 117
DB 3736 VLEDGNIKLTIQNAKEDVGARCEAVNVAGKANTNADIKIOPAKVHEHTVDESQOLE 3795
QY 118 V-QLVFGITANSDTHLQG-----QSLTLESPPGSSPSVQCR-----SPRGKNI 163
DB 3796 IQCFETVPGTASSKTDGTGGAPEFVELRSCTVTEK---QQAIIKCKYKGEPRPKIKWT 3851
QY 164 QGGKTLTSSV---QLELDSDSGTWTCTV--LQNKAKVEFKIDYPCAPAPKSKDCKT----- 213
DB 3851 KEGKEVENSARVRAHEKDDGTLITLPDNTQADAGEYR-----CEKENEYSAMTEGPIT 3906
QY 214 ---HTCPPELLGGPSVYLFPPPKPDITLMISRTPEVTCVVVDVSHEDPEVFENNYVDGVEYH 270
DB 3907 VTLEGAAPKIDGAPDLOPKPAVV-----TVGETAVLEGGKISCKPKPEVKKYKNGBELK 3961
QY 271 NAKTYPREEQVNSTYRVSVLTVLHQMINKGEYKCKVSKALPAPIEKTISKANGQPRE 330
DB 3962 PEDRAVKEIENLDGQFQ---LTVTNAKLDDMDYECEASNEFGDWSVTLT-----VKE 4012
QY 331 PQVYTLPPSRDELTKNOVSLT-----CLVKGFPYSDIAVEMESNGQPENNYKTPPVLD 384
DB 4013 P-AQVAPGFFKELSAIQVKEETAPKECVSGTKPD--VKWPKDGTPLKEDKRVAFEST 4068
QY 385 SDGSFFLYSKLTVDKSRW-QQGNVFSCSYMHEA 416
DB 4069 DDGT-----QRLVIEDSKTDDQGN-VRIEVSND 4096

```

```

RA Seling E., Miller J., Wilson R., Storb U.;
RT "Evolution of mouse immunoglobulin lambda genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:4681-4685(1982).
RN [2]
RA SEQUENCE OF 1-18 (MYELOMA PROTEIN CBPC-49).
RA Breyer R.M., Sauer R.T., Eisen H.N.;
RT "The variable region of mouse lambda-3 chains.";
RL ICN UCLA Symp. Mol. Cell. Biol. 20:105-110(1981).
RN [3]
RA SEQUENCE OF 10-104 (MYELOMA PROTEIN CBPC-49 AND MONOCLONAL ANTIBODY
RP 8-47).
RX MEDLINE=81223782; PubMed=6165398;
RA Azuma T., Steiner L.A., Eisen H.N.;
RT "Identification of a third type of lambda light chain in mouse
RL immunoglobulins.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:569-573(1981).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00585; AAB59670.1; -.
CC F01: B93922; L3MS.
CC DR HSP; P01842; 2MCG.
CC DR InterPro: IPR007110; IG-1like.
CC DR InterPro: IPR003597; IG-cl.
CC DR InterPro: IPR003006; IG_MHC.
CC DR Pfam: PF00047; Ig_1.
CC DR SMART; SM00407; IGc1; 1.
CC DR PROSITE; PS50835; IG_LIKE; 1.
CC DR PROSITE; PS00290; IG_MHC; 1.
CC KW Immunoglobulin domain; Immunoglobulin C region.
CC FT NON_TER 1 1
CC FT DOMAIN 6 99 IG-LIKE.
CC FT DISULFID 27 85
CC FT DISULFID 103 103 INTERCHAIN (WITH HEAVY CHAIN).
CC SQ SEQUENCE 104 AA; 11371 MW; 83CEBCD4AA348E1 CRC64;

Query Match
Best Local Similarity 34.3%; Score 163; DB 1; Length 104;
Matches 35; Conservative 16; Mismatches 47; Indels 4; Gaps 2;

QY 327 QPRE-POVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLD 385
DB 1 QPKSTPILTMPPSEBELQENKATLVCLISNPSGAVVANKANCTPITQGVDTSNPKYE 60
QY 386 DGSFFLYSKLTVDKSRWQGVNFSYVMEALHNYTKSLS 427
DB 61 DKWYASSFLHLTSPDQWRSHNSFTQVTHE---GDVTEKSL 99

```

```

RESULT 71
LAC3 MOUSE STANDARD; PRT; 104 AA.
ID LAC3 MOUSE STANDARD; PRT; 104 AA.
AC P01845;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE IG lambda-3 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83014953; PubMed=6812053;

```

```

RESULT 72
PGBM MOUSE STANDARD; PRT; 3707 AA.
ID PGBM MOUSE STANDARD; PRT; 3707 AA.
AC Q05753;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Melanoma;

```



```

FT DISULFID 344 359 BY SIMILARITY.
FT DISULFID 368 381 BY SIMILARITY.
FT DISULFID 375 394 BY SIMILARITY.
FT DISULFID 388 403 BY SIMILARITY.
FT DISULFID 428 479 BY SIMILARITY.
FT DISULFID 764 773 BY SIMILARITY.
FT DISULFID 766 780 BY SIMILARITY.
FT DISULFID 783 792 BY SIMILARITY.
FT DISULFID 795 811 BY SIMILARITY.
FT DISULFID 814 828 BY SIMILARITY.
FT DISULFID 816 839 BY SIMILARITY.
FT DISULFID 842 851 BY SIMILARITY.
FT DISULFID 854 869 BY SIMILARITY.
FT DISULFID 1159 1168 BY SIMILARITY.
FT DISULFID 1161 1175 BY SIMILARITY.
FT DISULFID 1178 1187 BY SIMILARITY.
FT DISULFID 1190 1206 BY SIMILARITY.
FT DISULFID 1209 1224 BY SIMILARITY.
FT DISULFID 1211 1234 BY SIMILARITY.
FT DISULFID 1237 1246 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1275 1287 BY SIMILARITY.
FT DISULFID 1277 1293 BY SIMILARITY.
FT DISULFID 1295 1304 BY SIMILARITY.
FT DISULFID 1307 1322 BY SIMILARITY.
FT DISULFID 1563 1572 BY SIMILARITY.
FT DISULFID 1565 1579 BY SIMILARITY.
FT DISULFID 1582 1591 BY SIMILARITY.
FT DISULFID 1594 1610 BY SIMILARITY.
FT DISULFID 1613 1628 BY SIMILARITY.
FT DISULFID 1615 1638 BY SIMILARITY.
FT DISULFID 1641 1650 BY SIMILARITY.
FT DISULFID 1653 1668 BY SIMILARITY.
FT DISULFID 1792 1839 BY SIMILARITY.
FT DISULFID 1886 1932 BY SIMILARITY.
FT DISULFID 1976 2021 BY SIMILARITY.
FT DISULFID 2073 2118 BY SIMILARITY.
FT DISULFID 2170 2215 BY SIMILARITY.
FT DISULFID 2266 2313 BY SIMILARITY.

```

Query Match Best Local Similarity 20.2%; Score 163; Length 3707; Pred. No. 0.033; Matchee 94; Conservative 66; Mismatch 199; Indels 106; Gaps 21;

```

QY 33 KQDTVELTTSAGOKSIOFHMKNSNOIKLNGOGSFLTGPKSKLNDRADSRSLMDQNF 92
DB 2065 ECGTDLNCGVPGHAAVOTMNRG-----GSLPTH-----HOTGSRRLY----- 2106
QY 93 PLIKNLKIEDSDTYICEVEDOKE--EVLVFGLTANSDTL-----L 134
DB 2107 -----QVSADSGEYCVSLSSGPLEASVLSITPAANVHITPGVVPPIRIETSSRYA 2161
QY 135 OGOSLTLLTLESPGSSPVQCRS-----PRKNIOGKTLVSOLELODSTWCTVLCN 189
DB 2162 ECGTDLNCGVPGHAAVOTMNRGSLPAGHGVH-GHMLRLNRVSPADSGEYSCVYCS 2220
QY 190 QAKVEKTIIV-----PCPAPPKSCDKTHTCPBELIGPSVLPFPKPKDTLMISRTPEV 244
DB 2221 SGTLEASVLTTEASEPSPIPAP-----GLAOPVYIESSSSHLTEGQTVDL 2266
QY 245 TCVVVDVSHDEPEVKENYVVDGVEVNAKTPREBOYNTVVSULTVLDHODMUNGKRY 304
DB 2267 KCVLPQQAHL-----AQTTMNRG-----SGLPAHQHGSLLRLYOLSPA-----DSGEY 2311
QY 305 KCKVSNKALP-----APIEKTISAKKGP-----REPQVTLTPPSRDELTKNOVSLTCLV-KGF 357
DB 2312 VCOVAASSHREHSAFKLTVPSQNSFRLSRVSISIEPPSSVVOGQGDASFCLIHGEA 2371
QY 358 YPSDLAVENESNGQP-ENNYKTPPLVDSGSEFLY---SKLTVDKSRMOGQNVFSC--S 411
DB 2372 MP--IKVEKIRNDQLEEDVNHISF-----NGSITTVAGPPTMEPTACVASNVMGMAQS 2424
QY 412 VNHVALHNHYTQKSLSLSP-----GLQDLBTCAEADGDELQGLWT 451

```

```

DB 2425 VNLNVHGPPTVSVLPBEGVHVHVKMDITLBEICISGGERSSPRWT 2469
|:::|
RESULT 73
ID IL1 HUMAN STANDARD; PRT; 213 AA.
AC P15814;
DT 01-APR-1990 (rel. 14, Created)
DT 01-APR-1990 (rel. 14, Last sequence update)
DT 15-MAR-2004 (rel. 43, Last annotation update)
DE Immunoglobulin lambda-like polypeptide 1 precursor (Immunoglobulin-
DE related 14.1 protein) (immunoglobulin omega polypeptide) (lambda 5)
DE (CD179b antigen).
OS IGL1 OR IGL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89315835; PubMed=2501791;
RA Hollis G.F., Evans R.J., Stafford-Hollis J.M., Korameyer S.J.,
RA McKearn J.P.;
RT "Immunoglobulin lambda light-chain-related genes 14.1 and 16.1 are
RT expressed in pre-B cells and may encode the human immunoglobulin
RT omega light-chain protein."
RL Proc. Natl. Acad. Sci. U.S.A. 86:5552-5556(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE=91108327; PubMed=1703205;
RA Evans R.J., Hollis G.F.;
RT "Genomic structure of the human Ig lambda 1 gene suggests that it may
RT be expressed as an Ig lambda 14.1-like protein or as a canonical B
RT cell Ig lambda light chain: implications for Ig lambda gene
RT evolution."
RL J. Exp. Med. 173:305-311(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=92288257; PubMed=12477932;
RA Strusberg R.L., Felsing E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schele C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hasle P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stanclevon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C.J.,
RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.F., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBMIT: Associates non-covalently with VPEEB1.
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN PRE-B-CELLS AND A SPECIAL
CC B-CELL LINE (WHICH IS SURFACE IG NEGATIVE).
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC -1- DATABASE: NAME=PROW; NOTE=PROW 1:64-67(2000);
CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/696419174_g.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M27749; AAA36100.1; -.
DR EMBL; M34513; AAA36096.1; -.
DR EMBL; M34511; AAA36096.1; JOINED.
DR EMBL; M34512; AAA36096.1; JOINED.
DR EMBL; BC012293; AAH12293.1; -.
DR PIR; A33911; A33911.
DR HSSP; P01842; 7FAB.
DR Genew; HGNC:5870; IGLL1.
DR MIM; 146770; -.
DR GO; GO:0016020; C:membrane; NAS.
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; Igc1; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR Antigen; Signal; Immunoglobulin domain.
FT SIGNAL 1 37 POTENTIAL.
FT CHAIN 38 213 IMMUNOGLOBULIN LAMBDA-LIKE POLYPEPTIDE 1.
FT DOMAIN 97 108 J REGION (BY SIMILARITY TO LAMBDA
FT DOMAIN 109 213 LIGHT-CHAIN).
SQ SEQUENCE 213 AA; 22963 MW; 913A742B943C79 CRC64;
Query Match
Best Local Similarity 30.5%; Score 162.5; DB 1; Length 213;
Matches 36; Conservative 21; Mismatches 44; Indels 17; Gaps 3;
QY 327 QPR-EPQVYTLPPSRDELTKQVSLTCLVKGFPSPDIWVESNGQP-ENNYKTPPYLD 384
DB 109 QPKATPSVTLFPSSSEELQANKATVLCIMNDPFGILVTWKADDTPTQGVEMTTPSKQ 168
QY 385 SDGSFFLYSKLTVDSKRWQGNVFCSSVMEALHNHYTKSLSPGLQDDETCAEQ 442
DB 169 SNNKYAASSYSLTLPEDQWRSSRSYSCQVME-----GSTVEKTAAPAE 211
RESULT 74
LAC2_MOUSE STANDARD; PRT; 104 AA.
AC P01844;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda-2 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83014953; PubMed=6812053;
RA Seising E., Miller J., Wilson R., Storb U.;
RT "Evolution of mouse immunoglobulin lambda genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:4681-4685(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82274221; PubMed=6287422;
RA Wu G.E., Govindji N., Hozumi N., Murialdo H.;
RT "Nucleotide sequence of a chromosomal rearranged lambda 2
RL Immunoglobulin gene of mouse.";
RN Nucleic Acids Res. 10:3831-3843(1982).
RN [3]
RP SEQUENCE FROM N.A. (MOPC 315).
RX MEDLINE=82220143; PubMed=6283385;

```

```

RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Somatic variants of murine immunoglobulin lambda light chains.";
RL Nature 298:380-382(1982).
RN [4]
RP SEQUENCE (MOPC 315).
RX MEDLINE=74048693; PubMed=4760498;
RA Dugan E.S., Bradshaw R.A., Sims E.S., Eisen H.N.;
RT "Amino acid sequence of the light chain of a mouse myeloma protein
RL (MOPC-315)."
RL Biochemistry 12:5400-5416(1973).
RN [5]
RP SEQUENCE OF 66-104 (MOPC 315), AND REVISIONS.
RX MEDLINE=81223782; PubMed=616598;
RA Azuma T., Steiner L.A., Eisen H.N.;
RT "Identification of a third type of lambda light chain in mouse
RL immunoglobulins.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:569-573(1981).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00595; AAA39151.1; -.
DR PIR; C93922; LZMS.
DR HSSP; P01842; 2MCG.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; Igc1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON TER 1 1
FT DOMAIN 6 99 IG-LIKE.
FT DISULFID 27 85
FT DISULFID 103 103 INTERCHAIN (WITH HEAVY CHAIN).
SQ SEQUENCE 104 AA; 11254 MW; CE4B67B868862D3 CRC64;
Query Match
Best Local Similarity 34.3%; Score 161; DB 1; Length 104;
Matches 35; Conservative 15; Mismatches 48; Indels 4; Gaps 2;
QY 327 QPRE-POVYTLPPSRDELTKQVSLTCLVKGFPSPDIWVESNGQPENNYKTPPYLD 385
DB 1 QPKATPSVTLFPSSSEELQANKATVLCISNFSGGTVAMKANGTITQGVDSINPTKE 60
QY 386 DGSFFLYSKLTVDSKRWQGNVFCSSVMEALHNHYTKSL 427
DB 61 GNKFMASFLHLSQWRSHNSFTQVTHE---GDVTEKSL 99
RESULT 75
HB2D_PIG STANDARD; PRT; 258 AA.
ID HB2D_PIG
AC P15983;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SIA Class II histocompatibility antigen, DQ haplotype D beta chain
DE precursor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.

```

RA MEDLINE=90361905; PubMed=2391424;  
 RA Guettersen K., Leguenn C., Hirsch F., Germana S., Pratt K.,  
 RA Sachs D.H.;  
 RT "Class II genes of miniature swine. IV. Characterization and  
 RT expression of two allelic class II DQB cDNA clones.";  
 RL J. Immunol. 145:1946-1951(1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 DR EMBL, M31498; AAA1085.1; --  
 DR HSP; P13760; 2589.  
 DR InterPro; IPR007110; Ig\_1like.  
 DR InterPro; IPR003597; Ig\_C1.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR000353; MHC\_II\_beta.  
 DR Pfam; PF00047; Ig\_1.  
 DR Pfam; PF00969; MHC\_II\_beta; 1.  
 DR ProDom; PD000328; MHC\_II\_beta; 1.  
 DR SMART; SM00407; IgC1; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR PROSITE; PS00280; IG\_MHC; 1.  
 DR MHC II; Transmembrane; Glycoprotein; Signal.  
 KM MHC II; Transmembrane; Glycoprotein; Signal.  
 FT CHAIN 1 31  
 FT SIGNAL 32 258  
 FT DOMAIN 32 123 SLA CLASS II HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 124 217 DQ HAPLOTYPE D BETA CHAIN.  
 FT DOMAIN 218 227 EXTRACELLULAR BETA-1.  
 FT TRANSMEM 228 248 CONNECTING PEPTIDE.  
 FT DOMAIN 249 258 CYTOPLASMIC TAIL.  
 FT DISULFID 44 108 BY SIMILARITY.  
 FT DISULFID 146 202 BY SIMILARITY.  
 FT CARBOHYD 48 48 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 258 AA; 29262 MW; E3AC75110AED47C3 CRC64;  
 Query Match 6.0%; Score 161; DB 1; Length 258;  
 Best Local Similarity 28.4%; Pred. No. 0.0015;  
 Matches 46; Conservative 28; Mismatches 80; Indels 8; Gaps 4;  
 QY 262 WYDGVENAKTKPREEQNSTYRVSVLTVLHODMLNGEKYCKVSKALPAPIEKTI 321  
 DB 54 WSVDRY-IYNOEEFLFSDDMGEYRAVTEPLGRPDADYLANGQKALEQKAEIDTVCKHNY 112  
 QY 322 SKAKG---OPREQVYTLPRSRDELTKQVSLTCLVKGFPYPDIAVENSQOPEN-NYK 377  
 DB 113 QIEEGTTLRRVQPTVITISPKAEALNHNLLVCAVTDYPPQVKVQFRNGQDEETAGV 172  
 QY 378 TTPPVLDSDGFPLYSKLTIVDKSRMOQGVNFCVGHALHN 419  
 DB 173 STPLIRNGD---WTGYVLVLMENLQRGDVTYTCRHSSLSQ 211  
 RESULT 76  
 VCAL HUMAN STANDARD; PRT; 739 AA.  
 ID VCAL HUMAN  
 AC P19320;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Vascular cell adhesion protein 1 precursor (V-CAM 1) (CD106 antigen)  
 DE (INCM-100).  
 GN VCAM1 OR L1CAM  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RX [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=umbilical vein;  
 RX MEDLINE=91016951; PubMed=1699207;  
 RA Pole T., Newman W., Gopal T.V.;  
 RT "Full length vascular cell adhesion molecule 1 (VCAM-1).";  
 RL Nucleic Acids Res. 18:5901-5901(1990).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90090619; PubMed=268898;  
 RA Osborn L., Hession C., Tizard R., Vassallo C., Luhnawsky S.,  
 RA Chi-Rosso G., Lobb R.;  
 RT "Direct expression cloning of vascular cell adhesion molecule 1, a  
 RT cytokine-induced endothelial protein that binds to lymphocytes.";  
 RL Cell 59:1203-1211(1989).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91352090; PubMed=1715583;  
 RA Cybulsky M.I., Fries J.W.U., Williams A.J., Sultan P., Eddy R.,  
 RA Byers M., Shows T., Gimbrone M.A. Jr., Collins T.;  
 RT "Gene structure, chromosomal location, and basis for alternative mRNA  
 RT splicing of the human VCAM1 gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7859-7863(1991).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91201302; PubMed=1707873;  
 RA Hession C., Tizard R., Vassallo C., Schiffer S.B., Gott D., Moy P.,  
 RA Chi-Rosso G., Luhnawsky S., Lobb R., Osborn L.;  
 RT "Cloning of an alternate form of vascular cell adhesion molecule-1  
 RT (VCAM1).";  
 RL J. Biol. Chem. 266:6682-6685(1991).  
 RN (5)  
 RP SEQUENCE FROM N.A., AND VARIANTS PHE-318; ALA-384; ALA-413 AND  
 RP LEU-716.  
 RA Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,  
 RA Rajkumar N.R., Toth E.J., Yi Q., Nickerson D.A.;  
 RT Submitted (Aug-2002) to the EMBL/Genbank/DBJ databases.  
 RN (6)  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RC TISSUE=Retinal pigment epithelium;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Kuzny D.M., Sodergren E.U., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield V.S.N., Krzywinski M.T., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN (7)  
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 25-226.  
 RX MEDLINE=95147978; PubMed=7531291;  
 RA Jones E.Y., Harlos K., Bottomley M.J., Robinson R.C., Driscoll P.C.,  
 RA Edwards R.M., Clements J.M., Dudgeon T.J., Stuart D.I.;  
 RT "Crystal structure of an integrin-binding fragment of vascular cell  
 RT adhesion molecule-1 at 1.8-A resolution.";  
 RL Nature 373:539-544(1995).  
 RN (8)  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.  
 RX MEDLINE=95296382; PubMed=7539295;  
 RA Wang J.-H., Pepinsky R.B., Stehle T., Liu J.-H., Karpusas M.,  
 RA Browning B., Osborn L.;



RT "The crystal structure of an N-terminal two-domain fragment of  
 RT vascular cell adhesion molecule 1 (VCAM-1): a cyclic peptide based on  
 RT the domain 1 C-D loop can inhibit VCAM-1-alpha 4 integrin  
 RT interaction.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 92:5714-5718(1995).  
 RM [9]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.  
 RA Wang J.-H., Stehle T., Peginsky R.B., Liu J.-H., Karpusas M.,  
 RA Osborn L.;  
 RA "Structure of a functional fragment of VCAM-1 refined at 1.9-A  
 RT resolution.";  
 RL Acta Crystallogr. D 52:369-379(1996).  
 CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION  
 CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1  
 CC INTEGRIN VIA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL  
 CC TRANSDUCTION. THE VCAM1/VIA4 INTERACTION MAY PLAY A  
 CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE  
 CC EMIGRATION TO SITES OF INFLAMMATION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=Long; IsoId=PI9320-1; Sequence=Displayed;  
 CC Name=Short; IsoId=PI9320-2; Sequence=VSP\_002580;  
 CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as  
 CC well as on macrophage-like and dendritic cell types in both normal  
 CC and inflamed tissue.  
 CC -1- INDUCTION: By cytokines (e.g. IL-1, TNF-alpha).  
 CC -1- PTM: Sialoglycoprotein.  
 CC -1- DISEASE: May play an important role in the genesis of  
 CC arteriosclerosis and rheumatoid arthritis.  
 CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.  
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD106 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd106.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; X53051; CAA37218.1; -;  
 DR EMBL; M30257; AAA51917.1; ALT\_TERM.  
 DR EMBL; M73255; AAA61270.1; -;  
 DR EMBL; M60335; AAA61269.1; -;  
 DR EMBL; AF536818; AAM96190.1; -;  
 DR EMBL; BC017276; AAM17276.1; -;  
 DR PIR; A41288; A41288.  
 DR PIR; B41288; B41288.  
 DR PDB; 1VCA; 15-SEP-95.  
 DR PDB; 1VSC; 20-JUN-96.  
 DR PDB; 1IJ9; 07-NOV-01.  
 DR Genew; HGNC:12663; VCAM1.  
 DR MIM; 192225; -;  
 DR InterPro; IPR003987; ICAM VCAM-1.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003989; VCAM-1.  
 DR Pfam; PF00047; Ig; 6.  
 DR PRINTS; PRO1472; ICAMVCAM1.  
 DR PRINTS; PRO1474; VCAM1.  
 DR SMART; SM00408; IGC2; 3.  
 DR PROSITE; PS50835; IG\_LIKE; 5.  
 KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane;  
 KW Repeat; Signal; Alternative splicing; Polymorphism; 3D-structure.  
 FT CHAIN 1 24  
 FT DOMAIN 25 739 VASCULAR CELL ADHESION PROTEIN 1.  
 FT TRANSMEM 699 720 EXTRACELLULAR (POTENTIAL).  
 FT

FT DOMAIN 721 739 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 25 105 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 109 212 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 223 309 IG-LIKE C2-TYPE 3.  
 FT DOMAIN 312 399 IG-LIKE C2-TYPE 4.  
 FT DOMAIN 408 506 IG-LIKE C2-TYPE 5.  
 FT DOMAIN 511 595 IG-LIKE C2-TYPE 6.  
 FT DOMAIN 600 684 IG-LIKE C2-TYPE 7.  
 FT DISULFID 47 95  
 FT DISULFID 52 99  
 FT DISULFID 137 195  
 FT CARBOHYD 273 273  
 FT CARBOHYD 365 365  
 FT CARBOHYD 417 417  
 FT CARBOHYD 463 463  
 FT CARBOHYD 531 531  
 FT CARBOHYD 561 561  
 FT VARSPLIC 310 402  
 FT VARIANT 318 318 S -> F.  
 FT VARIANT 384 384 /FTID=VAR\_014309.  
 FT VARIANT 413 413 T -> A.  
 FT VARIANT 413 413 /FTID=VAR\_014310.  
 FT VARIANT 716 716 G -> A.  
 FT VARIANT 716 716 /FTID=VAR\_014311.  
 FT VARIANT 716 716 I -> L.  
 FT STRAND 26 30 /FTID=VAR\_014312.  
 FT TURN 34 40  
 FT TURN 39 48  
 FT STRAND 43 50  
 FT STRAND 56 61  
 FT TURN 62 63  
 FT STRAND 70 74  
 FT TURN 75 76  
 FT STRAND 82 82  
 FT STRAND 87 89  
 FT HELIX 92 99  
 FT STRAND 100 101  
 FT TURN 102 114  
 FT STRAND 120 123  
 Query Match 5.9%; Score 160.5; DB 1; Length 739;  
 Best Local Similarity 19.5%; Pred. No. 0.0062;  
 Matches 114; Conservative 87; Mismatches 203; Indels 181; Gaps 25;  
 16 LALLPAAATGANKV-----VLGKGPTELTCTASOKKSIQPHWKNNOIKIG---NQ 65  
 15 LWMFPAASQAFKIETTPESRYLAQIGDSVSLTCTGCESPFFSMRTQIDSPILNGKVTNE 74  
 66 G--SFLTKGPSKLNDR-----ADSRRLMDQG-----NPP----- 93  
 75 GTTSTLTMTNPFVSFGNEHSYLTCTATCESRK--LEKGIQVEIYSFPKDEIHLSCPLEKGP 132  
 94 -----LIIKNLK-----TEDSD-----TYICEVED----- 113  
 133 ITVKCSVADVYPPDRLEIDILKGDHLMKSGQEPLEDDNRKSLFTKSLVETFTPVIEDIGKV 192  
 114 -----OKEEVOLLVFG-----LTNSDTHLLQCSLLTLESPPGS 149  
 193 LVCRAKLHIDEMDSVPTVROAVKELQYIISPKNTVIVSNPSTKLOEGSGVTMTCSSEGLP 252  
 150 SPSPV-----QCRSPRGKNIQGGKTLVSQLELDPSGTWTC-----TVLONOKKVEFKIDIVP 201  
 253 APEIPIWSKJDNNGNLHLSGNATVTLTAMMEDSGIYCGVNLIGKRRKVELIVQ--- 309  
 202 CPAPEPKSCDKTHTCCPELLGSPVFLPPPKPKTLMISRTPEVTCVVADVSHEDPEVKFN 261  
 310 -----EK-----PFTVEISPEGPRIAAQIGDSVMTLCSVMGC--ESPSPSWR 348  
 262 WYVDGVEVHNAKTKPREEOQNSYTRVVSVLTVTHQDMLNGEKYCKYS--NKALPAPIEK 319  
 349 TQIDS-----PLSGKVSSEGTNST-----LTLSPVSPFENHSYLTCTVCGHKLEKGIQV 398

```

OY 320 TISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPD-IAYE----- 365
DB 399 ELYSF---PRDEI---EMSGGLVNGSVTVSGKPSVYPLDLLEIELLKGTEILENIEF 452
OY 366 WESNGCPENNYKT-----TPVLSDSGSEFLYSKLTVDXSRMOQGVFSCSVMEALNNH 420
DB 453 LEDTDMKSLKNSLEMTFPTIEDTKALVCAKHLIDMEPEPPKRGSTQTLVYVAVR 512
OY 421 YTKSLIS---LSPGLQDLDETCADADGELGWTTPPRASALP 461
DB 513 DRTVLVSPSSILIEGSSVNMTCISQGFPAKILMSQLPENGELQF 557

RESULT 77
CAML BRARE
ID CAM1 BRARE STANDARD; PRT; 1197 AA.
AC 090478;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule LI.1 (N-CAM LI.1) (Fragment).
GN NADLI.1.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE=embryo;
RX MEDLINE=96155762; PubMed=8568941;
RT Tongiozci E., Bernhardt R.R., Schachner M.;
RT "Zebrafish neurons express two LI-related molecules during early
RT axonogenesis.";
RL J. Neurosci. Res. 42:547-561 (1995).
CC -1- FUNCTION: Cell adhesion molecule with an important role in the
CC development of the nervous system. Involved in neuron-neuron
CC adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
CC to axonin on neurons (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in postmitotic neurons in 16-36
CC hour embryos, including those in the brain, cranial ganglia and
CC olfactory placodes, and in all classes of spinal
CC neurons.
CC -1- DEVELOPMENTAL STAGE: Onset of expression correlates with the
CC initiation of axonogenesis in 16-36 hour embryos.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC LI/neurofascin/NG2CAM family.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X89204; CAA61490.1; -
CC FIR; T30581; T30581.
CC HSSP; P20241; 1CFB.
CC ZFIN; ZDB-GENE-980526-512; nadli.1.
CC InterPro; IPR008957; FN III-like.
CC InterPro; IPR003964; FN III.
CC InterPro; IPR003962; FNIII subd.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC Pfam; PF00041; fn3; 5.
CC PRINTS; PRO0014; FN3YPEIII.
CC SMART; SM00060; FN3; 5.

```

```

DR SMART; SM00408; IGC2; 4.
DR PROSITE; PS50835; IG_LIKE; 6.
KM Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;
KM Transmembrane; Repeac; Immunoglobulin domain.
FT NON_TER 1
FT DOMAIN <1 1054 EXTRACELLULAR (POTENTIAL).
FT TRAMEMEM 1055 1075 POTENTIAL.
FT DOMAIN 1076 1197 CYTOPLASMIC (POTENTIAL).
FT DOMAIN <1 58 IG-LIKE C2-TYPE 1.
FT DOMAIN 69 160 IG-LIKE C2-TYPE 2.
FT DOMAIN 165 263 IG-LIKE C2-TYPE 3.
FT DOMAIN 268 355 IG-LIKE C2-TYPE 4.
FT DOMAIN 360 442 IG-LIKE C2-TYPE 5.
FT DOMAIN 451 541 IG-LIKE C2-TYPE 6.
FT DOMAIN 546 638 FIBRONECTIN TYPE-III 1.
FT DOMAIN 645 739 FIBRONECTIN TYPE-III 2.
FT DOMAIN 744 849 FIBRONECTIN TYPE-III 3.
FT DOMAIN 850 948 FIBRONECTIN TYPE-III 4.
FT DOMAIN 952 1029 FIBRONECTIN TYPE-III 5.
FT DISULFID 92 143 BY SIMILARITY.
FT DISULFID 139 247 BY SIMILARITY.
FT DISULFID 289 339 BY SIMILARITY.
FT DISULFID 383 432 BY SIMILARITY.
FT DISULFID 472 525 BY SIMILARITY.
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 414 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1002 1002 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1027 1027 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1197 AA; 132860 MW; 7CE1505E9FC7B28 CRC64;

Query Match 5.9%; Score 160; DB 1; Length 1197;
Best Local Similarity 18.4%; Pred. No. 0.012;
Matches 107; Conservative 84; Mismatches 208; Indels 182; Gaps 25;

OY 9 HLLVLQALPLPAALQGNVVLGKGDYELTCTASQKKSIFPHKNSNOILNQSIF 68
DB 259 HYTVVEAA--PYWTRSPENHLVAPGETVRLDCKDGIAPAPITW--SINGVPSGTD--- 312
OY 69 LTKGPSKLNDRADRSRLMDQGNPLIINKLIEDSDTYICEVEDQKEVOLVGLGTAN 128
DB 313 -----VDPRRRV---SSGKLISNVFSPDTAVYQCAVANKGSILA-----N 351
OY 129 SDTHLLOGSLTLT-----LESPPGSSPSVQCS-----PRKNGIOGK 167
DB 352 THVHVVELPAQLTDERLYGATAGTWMLDCTGSPPLPKIHWEILDSIPLSNAKISQ 411
OY 168 T---LSVQLQLDQSGTTCVVLONQKVEKIDIVPCRAPSPKCDKTHTCPELGGP 223
DB 412 TTNGSKIKISNVSEEDSNRYTCSVSETNKSISADVEVL-----NNTK-----IVGP 456
OY 224 SVFLPPPKKDTLMTSRPEVTCVVVDVSHEDPEPKFMNVYGVGVHNAKTPREQVYS 283
DB 457 PQLHVIRSGDAIILCK-----YVDHNLKSPYQV--WNKGQ---HKTATSDNDKTHE 505
OY 284 TYRVVSVLTVLHQDWLNGKEYKCVNKAAPLIEKITSKAKGPREPOVYTLPPSRDEL 343
DB 506 IGSGLKVLVDQWQEDM---GIYSCVST--TLSDDTASGYITVQDKDPQSLKSEMER- 560
OY 344 TKNOVSLTCLVKGFPSPDIAVWESNGCPENNYKTTPLVD-----SDGSFPLYSKLT 396
DB 561 -----SVTISWMP--VENNSPTEYVIEWNEGETPDEGQWOKYRSVS 601
OY 397 VDKSRMOQGVFSCSVMEALH-----NHTOKSLIS----- 429

```

```

Db      602 QDIDSWR--SICSYKHFQIRAVNSIGTSAPTESSLSYSTPAKPDTPNENVTLTSDP 659
Qy      430 -----PGLDLDERCAEAGDELGWLTTPPPRSALPAP-----TGS 467
Db      660 KSMITISQMBMRORNGRPFQIKYFWRRAADS--GAHWI-----BSSVSNPPLMVNNTGT 712
Qy      468 -----ALPDPQTASALPDPAPASALPALAV 493
Db      713 FVSFEIKQAVNDLAAPEPLVIGYSGEDPFLKAPSLAV 753

RESULT 78
LAC_HUMAN ID STANDARD; PRT: 105 AA.
AC P01842; P80423;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ig lambda chain C regions.
GN IGLC1 AND IGLC2 AND IGLC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE (BENCE-JONES PROTEIN SH).
RX MEDLINE=70166723; PubMed=4909564;
RA Titani K., Wikler M., Shinoda T., Putnam F.W.;
RT "The amino acid sequence of a lambda type Bence-Jones protein. 3. The
RT complete amino acid sequence and the location of the disulfide
RT bridges."
RL J. Biol. Chem. 245:2171-2176(1970).
RN (2)
RP SEQUENCE (BENCE-JONES PROTEIN X), AND DISULFIDE BONDS.
RX MEDLINE=69088380; PubMed=4883841;
RA Milstein C., Clegg J.B., Jarvis J.M.;
RT "Immunoglobulin lambda-chains. The complete amino acid sequence of a
RT Bence-Jones protein."
RL Biochem. J. 110:631-652(1968).
RN (3)
RP SEQUENCE (BENCE-JONES PROTEIN NIG-64).
RX MEDLINE=83186114; PubMed=6404900;
RA Kametani F., Takayasu T., Suzuki S., Shinoda T., Okuyama T.,
RA Shindzu A.;
RT "Comparative studies on the structure of the light chains of human
RT immunoglobulins. IV. Assignment of a subsubgroup."
RL J. Biochem. 93:421-429(1983).
RN (4)
RP SEQUENCE (BENCE-JONES PROTEIN KERN).
RX MEDLINE=71150336; PubMed=5549568;
RA Poncetingl H., Hees M., Hilschmann N.;
RT "Structural rule of antibodies. Primary structure of a monoclonal
RT immunoglobulin-L-chain of the lambda type, subgroup IV (Bence-Jones-
RT protein Kern). V. The complete amino acid sequence and its genetic
RT interpretation."
RL Hoppe-Seyler's Z. Physiol. Chem. 352:247-266(1971).
RN (5)
RP PARTIAL SEQUENCE (MYELOMA PROTEIN NEMM).
RX MEDLINE=74109253; PubMed=4814727;
RA Chen B.L., Poljak R.J.;
RT "Amino acid sequence of the (lambda) light chain of a human myeloma
RT immunoglobulin (IgG New)."
RL Biochemistry 13:1295-1302(1974).
RN (6)
RP SEQUENCE (DOT).
RX MEDLINE=95255298; PubMed=7737190;
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-flavin monoclonal
RT immunoglobulins."
RL Eur. J. Biochem. 228:886-893(1995).
RN (7)
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF NEMM.
RX MEDLINE=75046825; PubMed=4215080;

```

```

RA Poljak R.J., Amzel L.M., Avey H.P., Chen B.L., Phizackerley R.P.,
RA Saul F.;
RT "The three-dimensional structure of the fab' fragment of a human
RT myeloma immunoglobulin at 2.0-A resolution."
RL Proc. Natl. Acad. Sci. U.S.A. 71:3440-3444(1974).
RN (8)
RP PARTIAL SEQUENCE (BENCE-JONES PROTEIN MCG).
RX MEDLINE=75013804; PubMed=4415202;
RA Fect J.W., Deutsch H.F.;
RT "Primary structure of the Mcg lambda chain."
RL Biochemistry 13:4102-4114(1974).
RN (9)
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF MCG.
RA Edmundson A.B., Ely K.R., Abola E.F., Schiffer M.,
RA Panagiotopoulos N.;
RT "Rational allotetramer and divergent evolution of domains in
RT immunoglobulin light chains."
RL Biochemistry 14:3953-3961(1975).
RN (10)
RP X-RAY CRYSTALLOGRAPHY OF MCG.
RX MEDLINE=90133913; PubMed=2515285;
RA Ely K.R., Herron J.N., Harker M., Edmundson A.B.;
RT "Three-dimensional structure of a light chain dimer crystallized in
RT water. Conformational flexibility of a molecule in two crystal
RT forms."
RL J. Mol. Biol. 210:601-615(1989).
RN (11)
RP SEQUENCE FROM N.A.
RX MEDLINE=82080680; PubMed=6273747;
RA Hiter P.A., Hoills G.P., Korsmeyer S.J., Waldmann T.A., Leder P.;
RT "Clustered arrangement of immunoglobulin lambda constant region genes
RT in man."
RL Nature 294:536-540(1981).
CC -I- MISCELLANEOUS: The sequence shown is the Kern-/Oz-/Mcg- chain
CC found in proteins SH, X, and NIG-64. The Kern protein has the
CC Kern+ marker, the NEMM protein has the Oz+ marker, the Mcg protein
CC has the Kern+ marker, and the Mcg+ marker.
CC -I- MISCELLANEOUS: Six tandem lambda-type genes were identified and
CC the 3 most 5' were sequenced. These correspond to the Mcg sequence
CC (lambda-1), the Kern-/Oz- sequence (lambda-2) and the Kern-/Oz+
CC sequence (lambda-3).
CC -I- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk.)
CC -----
DR EMBL: J00253; AA59107.1; -
DR EMBL: L38562; AB36581.1; ALT_INIT.
DR EMBL: X51754; CAB38569.1; ALT_INIT.
DR EMBL: X51755; CA36049.1; -
DR EMBL: X51755; CA36051.1; -
DR PIR: A92057; L2HU.
DR PDB: 2MCG; 15-JUL-92.
DR PDB: 7FAB; 31-JAN-94.
DR PDB: 1AOK; 04-FEB-98.
DR PDB: 1LIL; 15-MAY-97.
DR Genew: HGNC:5855; IGLC1.
DR Genew: HGNC:5856; IGLC2.
DR Genew: HGNC:5857; IGLC3.
DR MIM: 147220; -.
DR GO: GO:0003823; F:antigen binding; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003597; Ig cl.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: Pf00047; Ig; 1.
DR SMART: SM00407; IGL1; 1.
DR PROSITE: PS50835; IG_LIKE; 1.

```

```

DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Bence-Jones protein;
KW 3D-structure.
FT NON_TER 1 1
FT DOMAIN 6 100
FT DISULFID 27 86
FT DISULFID 104 104
FT VARIANT 5 5
FT VARIANT 7 7
FT VARIANT 7 7
FT VARIANT 45 45
FT VARIANT 56 56
FT VARIANT 82 82
FT STRAND 8 11
FT HELIX 15 19
FT TURN 20 21
FT STRAND 24 32
FT STRAND 38 43
FT TURN 44 45
FT STRAND 46 48
FT STRAND 52 59
FT TURN 61 62
FT STRAND 65 72
FT HELIX 75 80
FT STRAND 84 89
FT STRAND 94 99
SQ SEQUENCE 105 AA; 11236 MW; DCD9C7G201C13CC2 CRC64;

Query Match
Best Local Similarity 34.1%; Score 159; DB 1; Length 105;
Matches 31; Conservative 18; Mismatches 40; Indels 2; Gaps 2;

QY 327 OPR-EPQVYTLPPSRDELTKNOVSLCLYKGFPSDIAVEMSNQOP-ENNYKTPPVLD 384
DB 1 QPKAPSVTLFPSSSELDQNKATLVCLSDPFGAVTAMKADSSPVKAGVETTPSKQ 60
QY 385 SDGSPFLVSKLTIVDKSRMOQGNVFCGCVNHE 415
DB 61 SNMKVAASSYLSLTPEQWMSHRSSYSCQVTHE 91

RESULT 79
SHS1_HUMAN STANDARD; PRT; 503 AA.
ID SHS1_HUMAN STANDARD; Q8N517; Q8N517; Q8N517; Q9UDX2; Q9UDX2;
AC Q9YU9; 000683; 043799; Q8N517; Q8N517; Q8N517; Q9UDX2; Q9UDX2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase non-receptor type subfamily 1 precursor
DE (SHP substrate-1) (SHP-1) (Inhibitory receptor SHP-1) (Signal-
DE regulatory protein alpha-1) (SHP-alpha-1) (SHP-alpha-2) (SHP-alpha-
DE 3) (MYD-1 antigen) (Brain Ig-like molecule with tyrosine-based
DE activation motifs) (Blt) (Macrophage fusion receptor) (p84).
GN PTPN21 OR SHP1 OR SIRP OR MYD1 OR BIR OR MFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
RX MEDLINE=97223399; PubMed=90702220;
RA Yano T., Matcokaki T., Amato K., Matsuda Y., Takahashi N., Ochi F.,
RA Fujioaka Y., Kasuga M.;
RA "Mouse and human SHP-1: molecular cloning of cDNAs and chromosomal
RT localization of genes.";
RL Biochem. Biophys. Res. Commun. 231:61-67(1997).
RN [2]

```

```

RP SEQUENCE FROM N.A. (ISOFORM 1), POLYMORPHISM, PHOSPHORYLATION,
RP N-GLYCOSYLATION, AND INTERACTIONS WITH PTPN11, PTPN6 AND GRB2.
RC TISSUE=Placenta;
RX MEDLINE=97215901; PubMed=9062191;
RA Khalitonenkov A., Chen Z., Sures I., Wang H., Schilling J.,
RA Ullrich A.;
RT "A family of proteins that inhibit signalling through tyrosine kinase
RT receptors.";
RL Nature 386:181-186(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 3), AND VARIANTS ASP-95; LEU-96; ASN-100;
RP ARG-107; GLY-109 AND VAL-131.
RC TISSUE=Monocytes;
RX MEDLINE=98143722; PubMed=9485180;
RA Brooke G.P., Parsons K.R., Howard C.J.;
RT "Cloning of two members of the SIRP alpha family of protein tyrosine
RT phosphatase binding proteins in cattle that are expressed on monocytes
RT and a subpopulation of dendritic cells and which mediate binding to
RT CD4 T cells.";
RL Eur. J. Immunol. 28:1-11(1998).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS LEU-44; THR-50; THR-52;
RP ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107; GLY-109 AND
RP VAL-131.
RC TISSUE=Brain;
RX MEDLINE=20053880; PubMed=10585853;
RA Sano S.-I., Ohnishi H., Kubota M.;
RT "Gene structure of mouse BIR/SHP-1.";
RL Biochem. J. 344:667-675(1999).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ahearn J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,
RA Coulson A.G., Coville G.J., Dearden R., Dhoni P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharasaho M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., Murray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prichardlingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.B., Sehra H.K., Showkeen R., Sims S.,
RA Skue C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A.R., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilting L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS LEU-44; THR-50;
RP THR-52; ARG-54; ALA-57; GLY-75; ASP-95; LEU-96; ASN-100; ARG-107;
RP GLY-109 AND VAL-131.
RC TISSUE=Brain, Kidney, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Altshuler R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cabant T.M., Scheetz T.E.,
RA Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,

```

RA Bosak S., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodegren E.J., Lu X., Gibbs R.A.,  
 RA Whaley J., Helton E., Kettelman M., Maden A., Shvchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywnicki M.I., Skalka U., Smalins D.E.,  
 RA Butcher A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [7]  
 RN FUNCTION, AND INTERACTIONS WITH FVB; SCAP2 AND PTK2B.  
 RP MEDLINE=99401000; PubMed=10469599;  
 RX Timsa J.F., Swanson K.D., Mate-Cardine A., Raab M., Rudd C.E.,  
 RA Schraven B., Neel B.G.,  
 RA "SPS-1 is a scaffold for assembling distinct adhesion-regulated  
 RT multi-protein complexes in macrophages.";  
 RL Curr. Biol. 9:927-930(1999).  
 [8]  
 RN PHOSPHORYLATION BY JAK2, AND INTERACTIONS WITH PTPN11 AND JAK2.  
 RP MEDLINE=20428742; PubMed=10842184;  
 RA Stofega M.R., Aggeler L.S., Wang H., Ullrich A., Carver-Su C.,  
 RT "Negative regulation of growth hormone receptor/JAK2 signaling by  
 RT signal regulatory protein alpha.";  
 RL J. Biol. Chem. 275:28222-28229(2000).  
 [9]  
 RN FUNCTION, AND INTERACTION WITH CD47.  
 RP MEDLINE=21400825; PubMed=11509554;  
 RX Lacour S., Tanaka H., Demeure C., Mateo V., Rubio M., Brown E.U.,  
 RA Maliszewski C., Lindberg F.P., Oldenborg A., Ullrich A.,  
 RA Delvespesse G., Sarfati M.,  
 RT "Bidirectional negative regulation of human T and dendritic cells by  
 RT CD47 and its cognate receptor signal-regulator protein-alpha:  
 RT cell activation of IL-12 responsiveness and inhibition of dendritic  
 RT cell activation.";  
 RL J. Immunol. 167:2547-2554(2001).  
 [10]  
 RN -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts  
 CC as docking protein and induces translocation of PTPN6, PTPN11 and  
 CC other binding partners from the cytosol to the plasma membrane.  
 CC Supports adhesion of cerebellar neurons, neurite outgrowth and  
 CC glial cell attachment. May play a key role in intracellular  
 CC signaling during synaptogenesis and in synaptic function (By  
 CC similarity). Involved in the negative regulation of receptor  
 CC tyrosine kinase-coupled cellular responses induced by cell  
 CC adhesion, growth factors or insulin. Mediates negative regulation  
 CC of phagocytosis, mast cell activation and dendritic cell  
 CC activation. CD47 binding prevents maturation of immature dendritic  
 CC cells and inhibits cytokine production by mature dendritic cells.  
 CC -1- SUBUNIT: Binds PTPN11 when tyrosine-phosphorylated, except in  
 CC macrophages, where it primarily binds PTPN6. Binds GNB2 in vitro.  
 CC Binds FGR (By similarity). Binds JAK2 irrespective of its  
 CC phosphorylation status and forms a stable complex. Binds SCAP1  
 CC and/or SCAP2. The resulting complex recruits FVB. Binds PTK2B.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=P78324-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P78324-2; Sequence=VSP\_007030;  
 CC Note=No experimental confirmation available;  
 CC Name=3;  
 CC IsoId=P78324-3; Sequence=VSP\_007029;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in brain.  
 CC Detected on myeloid cells, but not T cells. Detected at lower  
 CC levels in heart, placenta, lung, testis, ovary, colon, liver,  
 CC small intestine, prostate, spleen, kidney, skeletal muscle and  
 CC pancreas.  
 CC -1- PTM: N-glycosylated.  
 CC -1- PTM: Phosphorylated on tyrosine residues in response to

```

CC stimulation with EGF, growth hormone, insulin and PDGF.
CC Dephosphorylated by PTPN11.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation in
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86043; BAA12974.1; -.
DR EMBL; Y10375; CAA71403.1; -.
DR EMBL; Y11047; CAA71944.1; ALT_INT.
DR EMBL; AB023450; BAA87929.1; -.
DR EMBL; AC004832; AAF19260.1; -.
DR EMBL; AL034562; CAB38874.1; -.
DR EMBL; AL049634; CAB84662.1; ALT_SEQ.
DR EMBL; AL117315; CAC12723.1; -.
DR EMBL; BC026682; AAB26682.1; -.
DR EMBL; BC033092; AAB33092.1; -.
DR EMBL; BC038510; AAC38510.1; -.
DR PIR; JC5287; JCS5287.
DR Genev; HGNC; 9662; PTPN11.
DR MIM; 602461; -.
DR GO; GO:0005886; C:plasma membrane; TAS.

```

Query Match	5.8k; Score 156.5; DB 1; Length 503;
Best Local Similarity 24.8k; Pred. No. 0.0069;	
Matches 75; Conservative 45; Mismatches 153; Indels 41; Gaps 14	
QY	136 GOSLTTLLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTVL-QNQKVE 194
Db	73 GPARELIYNQKEGHFPRVTTVSESTKRENNDFISISINITPADGATYCVAFRKQSPDTE 132
QY	195 FKIDIVCPAPPEPSCOKHTCTCEPLDGSVFLPPRPKDTLMTSRTEFLVCVVVDVSGE 254
Db	133 FK-----SGAGTELSTVAKPSAP-VVSGPAAKATP-----QHTVSFTCESHGFSR 177
QY	255 DPEVKFNMYVDGVEVHNAKTK--PREE---QYNSTYRVVSVLTVLHODMLNGKEYKCKV 308
Db	178 D--TTLKMFKNGNELSDFTNVDVGESVSYSISTAKVLTREDVHSQV-----CEV 229
QY	309 SNKALPA-PIEKT--ISKAKGQPREPVYTLPPSRDELTKQVSLTCLYKGFPSDIIVE 365
Db	230 AHTVTLQGDPLRGTAANLSETRRVPPTLEV-IQQPRAE---NQAVTLCVAKFTFPORLTL 285
QY	366 WESNGQPENNYKTPPYLSDSGSEFLYSKLTVDKSRQOQGNVFGCSVMHE--ALAHNYT 422
Db	286 WLENGNYSRTEIATSTVTENKDDGTYNMMSWILLVNVSAHRDVKLTTCQVEHGDQPAVSKSHD 345
QY	423 QKSLSLSPGLQDDETCAE 440
Db	346 LK-VSAHPKQEGSNTAAE 362
RESULT 80	
LAC_PIG	STANDARD; PRT; 105 AA.
AC	POI846;
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	IG lambda chain C region.
OS	Sus scrofa (Pig).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX	NCBI_TaxId=9823;
RN	[1]
RP	SEQUENCE.
XX	MEDLINE=78000254; PubMed=409425;

```

RA Novotny J., Franek F., Margolies M.N., Haber E.;
RT "Amino acid sequence of normal (microheterogeneous) porcine
RL immunoglobulin lambda chains";
CC Biochemistry 16:3763-3772(1977).
CC -1- MISCELLANEOUS: This chain was obtained from a mixture of normal
CC immunoglobulins.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC PIR: A02129; L1PG.
DR HSSP; P01842; 2MCG.
DR InterPro; IPR007110; IG-1-like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; IG1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON TER 1
FT DOMAIN 1 100
FT DISULFID 2 86
FT DISULFID 104 104
SQ SEQUENCE 105 AA; 11003 MW; 3817AEBD747C396 CRC64;

Query Match
Best Local Similarity 5.8%; Score 156; DB 1; Length 105;
Matches 32; Conservative 15; Mismatches 42; Indels 2; Gaps 2;

OY 327 QPR-EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQP-ENNYKTPPVLD 384
DB 1 QPKAPPLVTLFPPSSEELTKNATLVCLISDFYPGAVTVMKAGTIVTQGVETTKPSKQ 60
OY 385 SDGSFELYSKLTVDSRNOQGVSFSCVWHE 415
DB 61 SNKTYASSYTLASADWKSSEGFCTQVTHE 91

RESULT 81
LAC5_MUSSP STANDARD; PRT; 105 AA.
AC P20765;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda-5 chain C region.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_Taxid=10096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88196070; PubMed=3129289;
RA Mamt F., Cazenave P.A., Kindt T.J.;
RT "Conservation of the immunoglobulin C lambda 5 gene in the Mus gene.";
RL EMO J. 7:117-122(1988).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright; it is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M35582; AAA39152.1; -.
DR HSSP; P01842; 2MCG.
DR InterPro; IPR007110; IG-1-like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; IG1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.

```

```

KW Immunoglobulin domain; Immunoglobulin C region.
FT NON TER 1
FT DOMAIN 1 100
FT DISULFID 2 86
FT DISULFID 104 104
SQ SEQUENCE 105 AA; 11674 MW; AAB417DF68471A17 CRC64;

Query Match
Best Local Similarity 5.7%; Score 153.5; DB 1; Length 105;
Matches 35; Conservative 20; Mismatches 43; Indels 5; Gaps 3;

OY 327 QPR-EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQP-ENNYKTPPVLD 384
DB 1 QPKSDPLVTLFPLSLKLNQANKVTIVCLVSEFPETGLVDMKVDGVPVLTQGVETTKPSKQ 60
OY 385 SDGSFELYSKLTVDSRNOQGVSFSCVWHEALHNHYTKSL 427
DB 61 TNKTYMVSSYTLTLISDQWMPHSRYSCTVTHE---GNTVEKSVS 100

RESULT 82
LAC_RABIT STANDARD; PRT; 105 AA.
AC P01847;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig lambda chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=82091105; PubMed=6797414;
RA Garcia I., Jaton J.-C.;
RT "The primary structure of the constant region of Basilea-rabbit
RT immunoglobulin lambda-chains.";
RL Biochem. J. 197:177-183(1981).
CC -1- MISCELLANEOUS: This lambda chain expresses the c7 allotypic
CC specificity.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
DR HSSP; P01842; 7PAB.
DR InterPro; IPR007110; IG-1-like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; IG1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; FALSE_NEG.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON TER 1
FT DOMAIN 1 100
FT DISULFID 2 86
FT DISULFID 104 104
SQ SEQUENCE 105 AA; 11484 MW; B427513272E863D CRC64;

Query Match
Best Local Similarity 5.7%; Score 153; DB 1; Length 105;
Matches 32; Conservative 17; Mismatches 45; Indels 4; Gaps 2;

OY 331 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQP-ENNYKTPPVLDSDGSF 389
DB 6 PSVILFPPSSEELTKNATLVCLISDFYPRVTKVMKADGNSVTGQVDTTQPSKOSNNKY 65
OY 390 FLYSKLTVDKSRNOQGVSFSCVWHEALHNHYTKSL 427
DB 66 AASSFLHLTAQWKSYSQVTCQVTHE---GHTVEKSLA 100

RESULT 83
KACB_RABIT

```

```

ID KACB_RABIT STANDARD; PRT; 106 AA.
AC P01839;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE 19_Kappa-B4 chain C region.
GN K-BAS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Basilea;
RA Heidmann O., Roujeon F.;
RT "Multiplicity of constant kappa light chain genes in the rabbit
RL genome: a b4b4 homozygous rabbit contains a kappa-bas gene.";
EMBL J. 2:437-441 (1983).
CC -1- MISCLENEOUS: In Basilea rabbits, the major type of light chain
CC is lambda. The kappa chain shown is a minor component. All other
CC rabbit B allotypes have Cys-64.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; V01241; CAA24558.1; -
DR EMBL; V00885; -; NOT_ANNOTATED_CDS.
DR PIR; A02121; K4RBBS.
DR HSSP; P01842; 72AB.
DR InterPro; IPR007110; Ig-1like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; FALSE NEG.
DR Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 99 IG-LIKE.
FT DISULFID 27 87
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
SQ SEQUENCE 106 AA; 11279 MW; AF9B928DDA853849 CRC64;
Query Match 5.7%; Score 153; DB 1; Length 106;
Best Local Similarity 33.7%; Pred. No. 0.0016;
Matches 31; Conservative 19; Mismatches 36; Indels 6; Gaps 2;
Oy 326 GQPEPPQYVTLPPSRDELTKQVSLTCLVKCFYPSDIAVEMESGQPE---NNYKTPP 381
Db 1 GDPAPPSVLTLPPEPSKEELTGATATVCVANKFYPSDITVTKVGTQQSGIENSKT--P 58
Oy 382 VLDSGSEFLYSKLTVDKSRWQGNVPSCSYM 413
Db 59 QSPEDNTYLSSTLSTLSAQYNHSHSVYCEVV 90

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=90258861; Pubmed=2342461;
RA Schreie H., Thompson J., Bona M., Hefta L.J.F., Maruya A.,
RA Haasener M., Shively J.E., von Kleist S., Zimmermann W.;
RT "Cloning of the complete gene for carcinoembryonic antigen: analysis
RT of its promoter indicates a region conveying cell type-specific
RT expression.";
RL Mol. Cell. Biol. 10:2738-2748 (1990).
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=86038876; Pubmed=3670312;
RA Beauchemin N., Benchimol S., Cournoyer D., Fuks A., Stanners C.P.;
RT "Isolation and characterization of full-length functional cDNA clones
RT for human carcinoembryonic antigen.";
RL Mol. Cell. Biol. 7:3221-3230 (1987).
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=89122014; Pubmed=3220478;
RA Barnett T., Goebel S.J., Notdurft M.A., Elting J.J.;
RT "Carcinoembryonic antigen family: characterization of cDNAs coding
RT for NCA and CEA and suggestion of nonrandom sequence variation in
RT their conserved loop-domains.";
RL Genomics 3:59-66 (1988).
RN
RP SEQUENCE OF 5-702 FROM N.A.
RX MEDLINE=87128144; Pubmed=3814146;
RA Okawa S., Nakazato H., Kosaki G.;
RT "Primary structure of human carcinoembryonic antigen (CEA) deduced
RT from cDNA sequence.";
RL Biochem. Biophys. Res. Commun. 142:511-518 (1987).
RN
RP SEQUENCE OF 331-702 FROM N.A.
RX MEDLINE=87204247; Pubmed=3033671;
RA Zimmermann W., Ortlieb B., Friedrich R., von Kleist S.;
RT "Isolation and characterization of cDNA clones encoding the human
RT carcinoembryonic antigen reveal a highly conserved repeating
RT structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2960-2964 (1987).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: Found in adenocarcinomas of endoderally
CC derived digestive system epithelium and fetal colon.
CC -1- PTM: COMPLEX IMMUNOREACTIVE GLYCOPROTEIN WITH A MW OF 180 KDA
CC COMPRISING 60% CARBOHYDRATE.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. CEA family.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like domains.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD66e entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd66e.htm.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M17303; AAB59513.1; -
DR EMBL; M59262; AAA62835.1; ALT SEQ.
DR EMBL; M59255; AAA62835.1; JOINED.
DR EMBL; M59256; AAA62835.1; JOINED.
DR EMBL; M59257; AAA62835.1; JOINED.
DR EMBL; M59258; AAA62835.1; JOINED.
DR EMBL; M59259; AAA62835.1; JOINED.
DR EMBL; M59260; AAA62835.1; JOINED.
DR EMBL; M59261; AAA62835.1; JOINED.
DR EMBL; M59709; -; NOT_ANNOTATED_CDS.
DR EMBL; M59710; -; NOT_ANNOTATED_CDS.
DR EMBL; M29540; AAA51967.1; -

```



[illegible][illegible]

```

RESULT 65
KACA_RAT
ID_KACA_RAT STANDARD: PRT: 106 AA.
AC P01836;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig kappa chain C region, A allele.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA;
RX MEDLINE=82082587; PubMed=6273908;
RA Sheppard H.W., Gutman G.A.;
RT "Allelic forms of rat kappa chain genes: evidence for strong
RT selection at the level of nucleotide sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:7064-7068(1981).
CC -I- SIMILARITY: Contains 1 immunoglobulin-like domain.
DR PIR, A02118; K1RRA.
DR HSBP, P01842; 2MCG.
DR InterPro; IPR007110; Ig-1like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00407; IgC1; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region.
KW NON TER 1
FT DOMAIN 5 102 IG-LIKE.
FT DISULFID 26 86
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
SQ SEQUENCE 106 AA; 11732 MW; B7E120D9700DD66 CRC64;

Query Match 5.6%; Score 150; DB 1; Length 106;
Best Local Similarity 31.5%; Pred.No. 0.0025;
Matches 29; Conservative 23; Mismatches 28; Indels 12; Gaps 2

331 PQQVTLPSRDELTKNGVSLTCLVKGFPSPDIAYEMWESGQPENNYKTTPTPLVDS----- 385
Db 5 PTVSIFPSMQLVSGGATVVCVNNFFPRDISVYKMKIDGSEQRD-----GLVDSVTQDD 59

```

```

Oy 386 -DCGFFPKLTVKDSMOGNGVSCGVMH 415
Db 60 SKDSTYSMSSTLSTKTYEYERHNLTYCEVHK 91

RESULT 86
HB2C PIG ID HB2C PIG STANDARD; PRt; 261 AA.
AC P159B2;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SLA class II histocompatibility antigen, DQ haplotype C beta chain
precursor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
ON NCBI_TaxID=9623;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=90361905; PubMed=2391424;
RA Gustafsson K., Leguenn C., Hirsch F., Germans S., Pratt K.,
RA Sachs D.H.;
RA "Class II genes of miniature swine. IV. Characterization and
expression of two allelic class II DQB CDNA clones.";
RL J. Immunol. 145:1946-1951(1990).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M31497; AAA531084.1; -
DR EMBL; M32117; AAA53110.1; -
DR PIR; A60404; A60404.
DR HSSP; P13760; 2SEB.
DR InterPro; IPR007110; IG-1like.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00047; Ig_1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGc1; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC II; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 31
FT CHAIN 32 261
FT FT
FT DOMAIN 32 126 SLA CLASS II HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 127 220 DQ HAPLOTYPE C BETA 1.
FT DOMAIN 221 230 EXTRACELLULAR BETA-1.
FT TRANSMEM 231 251 EXTRACELLULAR BETA-2.
FT DOMAIN 252 261 CONNECTING PEPTIDE.
FT DISULFID 47 111 CYTOPLASMIC TAIL.
FT DISULFID 149 205 BY SIMILARITY.
FT CARBOHYD 51 51 N-LINKED (GLCNAC...)(POTENTIAL).
SO SEQUENCE 261 AA; 29550 MW; AA9581F2A3B1969D CRC64;

Query Match 5.5%; Score 149.5; DB 1; Length 261;
Best Local Similarity 23.2%; Pred. No. 0.0085;
Matches 56; Conservative 40; Mismatches 106; Indels 39; Gaps 8;

Oy 180 GTWCTCTVQNGKQKVEFKIDIVPCRPBPKSCDKHTTCGLGPEVFLPPPKQTLWIS 239
Db 12 GLMTAL-----TMTLVVLGAPVAEGRD-----SPDVFQPK-GECTYYN 51
Oy 240 RTPEYTCVVDVSHEDPEVKENWYDGVVNAKTKPREEQYNSTYRVASVLYLADQML 299

```

[illegible]





```

FT CARBOHYD 853 853 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 994 994 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1009 1009 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1133 1133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1150 1150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1171 1171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 987 993 /FTid=VSP 008935.
/FTid=VSP 008935.
FT VARSPLIC 1132 1222 Missing (in isoform 2).
/FTid=VSP 008936.
FT CONFLICT 1105 1105 T -> A (in REF. 2).
SQ SEQUENCE 1369 AA; 152954 MW; 835F27D086B2BF6 CRC64;
Query Match 5.5%; Score 148; DB 1; Length 1369;
Best Local Similarity 22.3%; Pred. No. 0.085;
Matches 107; Conservative 63; Mismatches 182; Indels 128; Gaps 31;
QY 4 GVPF-----RHLLVLQAL-----LP-----AATQGNKVVVGKKGVTELTCTASQKS 48
DB 12 GIAFLALCLHLLISAIEVPLDSNIQSELPQPTITQSVVDYIVDRDNIPIECBAKNPV 71
QY 49 IOFWKNSNQIKILQNGSF--LTGPKSLNDRADSRSLMD-----QGNFPLII 96
DB 72 PTFSWTRN-----GKFFNVAKDP-KVSWRRSGTLVIDFHGGRPDVEGEYCFA 121
QY 97 KN-----LKIEDSDYICEVDEQKEVOLLVGLTANSTHLLQGSLLTLLES 145
DB 122 RNDVGTALSSKIHLOVSRSPLM-----PKEKVDYI-----EVDGAPSLQCNP 165
QY 146 PGSSPSV-----QCRSP--RGKNIQGGKT--LSVSOLELQDSGT--WTCT-----VLQ 188
DB 166 PFGLEPPIVIFMSSMERPIHODKRVSOQNGDLVSNWMLQADQIDSCNARFHTHTIQ 225
QY 189 NOKKVEFKIDYPCAPPEPKSCDKHTTCELLG-----PSVLEPPPKDTLMSRPE 243
DB 226 OKNPTLTK---VTKKPKHNETSLRNHTDMYSARGTETTPS-FMYPGTSSSQWLRGVD 281
QY 244 --VTCVVVDVSHEDPEVKFNVVDGVEVNAKTRREGQNSTYRVSVLVTLQHMDLNG 301
DB 282 LLLECTIAGV-----PAPDPMYKKGELPAGKTK--LENFNALNISNVEE-----DS 329
QY 302 KEYKCVSNKALPARIKTIIS-KAKGP---REPVTLLPSPRDELTKNOVSLTCLVKG 357
DB 330 GEYFLAANKM--GSIRHTISVRVKAAPYKWDPEPQMLIABEDG-----RLVBRAG- 380
QY 358 YPSDIAVESNGQENNYKTPPVLDSD--GSFLYSKLTVDKSRMOQGVFSCVWHE 415
DB 381 NPKP-SIQMLVNGEP---IEGSPFNPSPREVAGDTIVFRDTQIGSSA-----VYQCNAGNE 431
RESULT 90
LAMP HUMAN STANDARD; PRT; 338 AA.
AC Q13449;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
OS LSAMP OR LAMP.
OS Homo sapiens (Human).
OC Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96235133; PubMed=8666243;
RA Pimenta A.P., Fischer I., Levitt P.;
RT "cDNA cloning and structural analysis of the human limbic-system-
associated membrane protein (LAMP).";
RL Gene 170:189-195(1996).
CC -1- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF

```

```

CC CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: Expressed on limbic neurons and fiber tracts
CC as well as in single layers of the superior colliculus, spinal
CC chord and cerebellum.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U41901; AAC50569.1; -.
CC DR PIR; JCA4776; JCA4776.
CC DR Genew: HGNC:6705; LSAMP.
CC DR MIM; 603241; -.
CC DR GO; GO:0007399; P:neurogenesis; TAS.
CC DR InterPro; IPR007110; IG-1like.
CC DR InterPro; IPR003598; IG_c2.
CC DR Pfam; PF00047; Ig_3.
CC DR SMART; SM00408; IGC2; 2.
CC DR PROSITE; PS50835; IG_LIKE; 3.
CC KM Immunoglobulin domain; Cell adhesion; glycoprotein; GPI-anchor;
CC Repeat; Signal; Lipoprotein.
CC FT SIGNAL 1 28
CC FT CHAIN 29 315
CC FT PROPEP 316 338
CC FT DOMAIN 29 122
CC FT DOMAIN 132 214
CC FT DOMAIN 219 304
CC FT DISULFID 53 111
CC FT DISULFID 153 197
CC FT DISULFID 239 290
CC FT CARBOHYD 46 40
CC FT CARBOHYD 60 66
CC FT CARBOHYD 136 136
CC FT CARBOHYD 148 148
CC FT CARBOHYD 279 279
CC FT CARBOHYD 287 287
CC FT CARBOHYD 300 300
CC FT CARBOHYD 315 315
CC FT LIPID 315 315
CC SQ SEQUENCE 338 AA; 37308 MW; 0345F286DF5D92F CRC64;
Query Match 5.4%; Score 147; DB 1; Length 338;
Best Local Similarity 22.7%; Pred. No. 0.017;
Matches 75; Conservative 49; Mismatches 129; Indels 76; Gaps 16;
QY 10 LLVLVQLALPAA-----TQGNKVVLGKGGDTVELTCTASQKSIQFWKNSNQIKI 61
DB 14 LVLLRLCLLPGLPVRSVDNPGTNDITVYQGGTALIRCLVLEBKNS-KVALNRSGLIF 72
QY 62 LGNGSFLTKGPKSLNDA--DSRRSLMDQGNFPLINKLIEDSDYICEVDEQKE--- 116
DB 73 AGHD-----KMSLDPRVLEKHSI---EYSLRQKVVDVYBGSTCSVQTOHEKPT 121
QY 117 -EVOLLVFG-----LTANSDTHLQGSLLTLSPSSPSVQCR--SPRGKNIQGGKT- 168
DB 122 SQVTLIVGPPKINISSDVTVNNGSNVTLVCMANGREPIYTRHRLPFREREGBEY 181
QY 169 LSVSOLELQDSGTWTCVYLQ-----NOKKVEFKIDYPCAPPEPKSCDKHTTCELLG 223
DB 182 LEIGITREQSGKYECKANESVADYQVAVNYNYP-TTTEKSNENAT----- 230
QY 224 SVFLPPPKDTLMSRPEVTCVVVDVSHEDPEVKFNVVDGVEVNA-----KTKPREQ 280

```

Db 231 -----TGROASLKCEASAV-----PAPDEFWYDDPTRINSANGLERKSTEGQ 272

QY 281 YNSTYRVSVLTVLHQMVLNGKEYCKKYSNK 311

Db 273 SSLT-----YTNVTEHYGN---YTCVAANK 295

RESULT 91

NPAS\_MOUSE STANDARD; PRT; 1240 AA.

AC Q81003;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Neurofascin precursor.

GN NPASC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Swiss Webster; TISSUE=Brain;

RA Dicks P., Montag-Sallaz M., Montag D.;

RT "Expression patterns of Ll-family cell recognition molecules Ll, CHL1, NrcAM, and neurofascin in the mouse brain."

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Cell adhesion, ankyrin-binding protein which may be involved in neurite extension, axonal guidance, synaptogenesis, myelination and neuron-glia cell interactions (by similarity).

CC -1- SUBUNIT: Probable constituent of a neurofascin/NRCAM/ankyrin G complex. Associates with the sodium channel beta-1 (SCN1B) and beta-3 (SCN1B) subunits (by similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.

CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.

CC -1- SIMILARITY: Contains 4 fibronectin type III domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AJ543322; CAB5849.1; -

DR MCD; MG1:2442229; D430023G06R1k.

DR InterPro: IPR003961; FN.III.

DR InterPro: IPR008957; FN.III-like.

DR InterPro: IPR003599; IG.II-like.

DR InterPro: IPR007110; IG.II-like.

DR InterPro: IPR003598; IG.C2.

DR Pfam; PF00041; fn3; 4.

DR Pfam; PF00047; Ig; 6.

DR SMART; SMO0060; FN3; 4.

DR SMART; SMO0409; IG; 6.

DR SMART; SMO0408; IGC2; 6.

DR PROSITE; PS50835; IG LIKE; 6.

KM Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain; Glycoprotein.

FT SIGNAL 1 24

FT CHAIN 25 1240

FT DOMAIN 25 1110

FT TRANSSEM 1131 1131

FT DOMAIN 1132 1240

FT DOMAIN 41 137

FT DOMAIN 143 230

FT DOMAIN 244 332

FT DOMAIN 337 424

FT DOMAIN 430 517

POTENTIAL.

NEUROFASCIN.

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

IG-LIKE C2-TYPE 1.

IG-LIKE C2-TYPE 2.

IG-LIKE C2-TYPE 3.

IG-LIKE C2-TYPE 4.

IG-LIKE C2-TYPE 5.

FT DOMAIN 521 603

FT DOMAIN 628 720

FT DOMAIN 727 820

FT DOMAIN 825 924

FT DOMAIN 1008 1094

FT DOMAIN 924 1006

FT DISULFID 63 118

FT DISULFID 162 213

FT DISULFID 268 316

FT DISULFID 358 408

FT DISULFID 452 501

FT DISULFID 543 592

FT CARBOHYD 305 305

FT CARBOHYD 409 409

FT CARBOHYD 446 446

FT CARBOHYD 483 483

FT CARBOHYD 752 752

FT CARBOHYD 778 778

FT CARBOHYD 866 866

FT CARBOHYD 881 881

SO SEQUENCE 1240 AA; 137975 MW; 6DB8935B5B02E965 CRC64;

Query Match 5.4%; Score 147; DB 1; Length 1240;

Best Local Similarity 21.4%; Pred. No. 0.087;

Matches 100; Conservative 66; Mismatches 183; Indels 118; Gaps 26;

QY 6 PFRHLVLVQLALLPA-----TGCKNVVIGKKGTVELCTASQ 45

Db 8 PWHIALILFLSLGGHIEIPMDPSIONELTOPPTTKQSVKHIVDPNNILIECAKG 67

QY 46 KKSIOFHW-NNSNQIKILNGQGSFLTK--GPSKLNDRAUSRSIMQGNFPLIKN----- 98

Db 68 NPAPSPFWMTNSRFPNIAKOPRVSMRRSGTLVIDFSGGRPREY-BGEYQCFARNKFGT 126

QY 99 -----LKEDSDTYICEVEDQKEVQLVFGLTANSDFHLQGSLLTLESPPG-SS 150

Db 127 ALSNRIRLQVSKSPW-----PKENLDPVV-----VOGAPLTLQCNPPGLPS 170

QY 151 PSYQCSPPKKNIGGGKTLT-----VSQLELDQSGT-WTCGVION-----QKKEFK 196

Db 171 PVIFWSSSMIEPTIDQKRVSGHNGDLVYSNVMLQDQITYSCHARHFHTTIDQKPF 230

QY 197 IDIVCPAPPEPKSCDHTHTCPBELLGSPVLPFPKKDPTLMISRTPE--VTCVVVDVSH 254

Db 231 LKVLTRGVAERT-----PS-FMYQGTSSSQMVLRGMDDLIECIASGV--- 273

QY 255 DPEVKFNWYDGVVHNAAKTPREBOYNSYRVSVLTVLHQMVLNGKEYCKKYSNKALP 314

Db 274 -PTPDIAWYKKGDDLPSNKA--FENFNKALRTYNSEE-----DSGEYFCIASNKA-- 322

QY 315 APIEKTIS-KAKGP---REPOVYTLPPSRDELTKQVSLTCLVKGYPSDIAVEMESNG 370

Db 323 GSIRHTISRVAKAPVWLDPPKVLILAPGEDG-----RLVCRANG-NPAP-TVQMWVNG 374

QY 371 QPENNYKTPPVLDSD--GSFLLYSKLTVDKSRMOQGNFSCSVME 415

Db 375 EF---LQSAPEPNRREVAAGDTIIFRDTQISSRA-----VYQCNTSNE 413

RESULT 92

NPEN\_MOUSE STANDARD; PRT; 1242 AA.

AC Q9QZ57;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Nephlin precursor (Renal glomerulus-specific cell adhesion receptor).

GN NPEN1 OR NPEN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.

MEDLINE=994363468; PubMed=10504499;

Holtzman L.B., St John P.L., Kovari I.A., Verma R., Holtzoefer H.,

Abrahamson D.R.;

"Nephrin localizes to the slit pore of the glomerular epithelial

cell.";

Kidney Int. 56:1481-1491(1999).

[2]

INTERACTION WITH CD2AP.

MEDLINE=21590051; PubMed=11733379;

Shih N.Y., Li J., Cotran R., Mundel P., Miner J.H., Shaw A.S.;

"CD2AP localizes to the slit diaphragm and binds to nephrin via a

novel C-terminal domain.";

Am. J. Pathol. 159:2303-2308(2001).

[3]

INTERACTION WITH CD2AP AND NPHS2.

MEDLINE=21590460; PubMed=11733557;

Schwarz K., Simons M., Reiser J., Saleem M.A., Faul C., Kriz W.,

Shaw A.S., Holtzman L.B., Mundel P.;

"Podocin, a fatc-associated component of the glomerular slit

diaphragm, interacts with CD2AP and nephrin.";

J. Clin. Invest. 108:1621-1629(2001).

-1- FUNCTION: Seems to play a role in the development or function of

the kidney glomerular filtration barrier. May anchor the podocyte

slit diaphragm to the actin cytoskeleton.

-1- SUBUNIT: Interacts with podocin/NPHS2 and with CD2AP C-terminal

domain.

-1- SUBCELLULAR LOCATION: Type I membrane protein (Potential). Located

at podocyte slit diaphragm between podocyte foot processes.

-1- TISSUE SPECIFICITY: Expressed in kidney glomeruli.

-1- PTM: Phosphorylated on tyrosine residues (By similarity).

-1- SIMILARITY: Belongs to the immunoglobulin superfamily.

-1- SIMILARITY: Contains 8 immunoglobulin-like domains.

-1- SIMILARITY: Contains 1 fibronectin type III domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>

or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

-----

EMBL; AF168466; AAFO3368.1; -

MCD; MG1:1859637; Nphs1.

GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0007254; P:JNK cascade; IDA.

GO; GO:0000165; P:MAPKK cascade; IDA.

InterPro; IPRO08957; FN\_III-like.

InterPro; IPRO03961; FN\_III.

InterPro; IPRO07110; IG-like.

InterPro; IPRO03598; IG\_c2.

Pfam; PF00041; fn3; 1.

Pfam; PF00047; igf; 8.

SMART; SMO0060; FN3; 1.

SMART; SMO0408; IGc2; 1.

PROSITE; PS50835; IG LIKE; 8.

Cell adhesion; Transmembrane; Signal; Glycoprotein;

Immunoglobulin domain; Repeat; Phosphorylation.

FT SIGNAL 1 22

FT CHAIN 23 1242

FT DOMAIN 23 1066

FT TRANSMEM 1065 1086

FT DOMAIN 1087 1242

FT DOMAIN 25 130

FT DOMAIN 135 233

FT DOMAIN 242 333

FT DOMAIN 340 434

FT DOMAIN 440 540

FT DOMAIN 544 635

FT DOMAIN 740 832

FT DOMAIN 838 939

FT DOMAIN 941 1025

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

IG-LIKE C2-TYPE 1.

IG-LIKE C2-TYPE 2.

IG-LIKE C2-TYPE 3.

IG-LIKE C2-TYPE 4.

IG-LIKE C2-TYPE 5.

IG-LIKE C2-TYPE 6.

IG-LIKE C2-TYPE 7.

IG-LIKE C2-TYPE 8.

FIBRONECTIN TYPE-III.

FT	DISULFID	53	111	POTENTIAL.
FT <th>DISULFID</th> <td>160</td> <td>217</td> <td>POTENTIAL.</td>	DISULFID	160	217	POTENTIAL.
FT <th>DISULFID</th> <td>265</td> <td>317</td> <td>POTENTIAL.</td>	DISULFID	265	317	POTENTIAL.
FT <th>DISULFID</th> <td>361</td> <td>417</td> <td>POTENTIAL.</td>	DISULFID	361	417	POTENTIAL.
FT <th>DISULFID</th> <td>465</td> <td>528</td> <td>POTENTIAL.</td>	DISULFID	465	528	POTENTIAL.
FT <th>DISULFID</th> <td>567</td> <td>623</td> <td>POTENTIAL.</td>	DISULFID	567	623	POTENTIAL.
FT <th>DISULFID</th> <td>761</td> <td>816</td> <td>POTENTIAL.</td>	DISULFID	761	816	POTENTIAL.
FT <th>DISULFID</th> <td>863</td> <td>920</td> <td>POTENTIAL.</td>	DISULFID	863	920	POTENTIAL.
FT <th>CARBOHYD</th> <td>40</td> <td>40</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	40	40	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>356</td> <td>356</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	356	356	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>401</td> <td>401</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	401	401	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>547</td> <td>547</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	547	547	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>553</td> <td>553</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	553	553	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>564</td> <td>564</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	564	564	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>577</td> <td>577</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	577	577	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>680</td> <td>680</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	680	680	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>708</td> <td>708</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	708	708	N-LINKED (GLCNAC . . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <td>908</td> <td>908</td> <td>N-LINKED (GLCNAC . . .) (POTENTIAL.)</td>	CARBOHYD	908	908	N-LINKED (GLCNAC . . .) (POTENTIAL.)
SO <th>SEQUENCE</th> <td>1242</td> <td>AA, 134890</td> <td>MM, 02DB2180BF145092 CRC64;</td>	SEQUENCE	1242	AA, 134890	MM, 02DB2180BF145092 CRC64;

Query Match 5.4%; Score 147; DB 1; Length 1242;  
 Best Local Similarity 20.6%; Pred. No. 0.087;  
 Matches 104; Conservative 61; Mismatches 219; Indels 122; Gaps 21

QY	11	LVLTQALLPAPATG	-----	NKVYLGKKGDVETELCTAQSQKSIQFIHWKNSNQIKI	61
DB <td>16</td> <td>MLTTLGLASPVPTAPRPFMAISNLTVV</td> <td>-----</td> <td>EGSTIKLMCGVRAPGSV-VQM</td> <td>64</td>	16	MLTTLGLASPVPTAPRPFMAISNLTVV	-----	EGSTIKLMCGVRAPGSV-VQM	64
QY <td>62</td> <td>LGNGGSPLTKPS</td> <td>-----</td> <td>KLNDRADSRSLMDQGNFLIINKLIEDSDTYICEVEDQKE</td> <td>116</td>	62	LGNGGSPLTKPS	-----	KLNDRADSRSLMDQGNFLIINKLIEDSDTYICEVEDQKE	116
DB <td>65</td> <td>-AKDGLLGPNPKIPGPPRYSLEEDSA</td> <td>-----</td> <td>-GPHILITACDLSDAEYECQVGRSEL</td> <td>118</td>	65	-AKDGLLGPNPKIPGPPRYSLEEDSA	-----	-GPHILITACDLSDAEYECQVGRSEL	118
QY <td>117</td> <td>EVQLL</td> <td>-----</td> <td>VFGLT--ANSDTHLQGSITLTLESPPGSSPSVQCSPPRG</td> <td>161</td>	117	EVQLL	-----	VFGLT--ANSDTHLQGSITLTLESPPGSSPSVQCSPPRG	161
DB <td>119</td> <td>GPELVSRVILSVLVPRKVLQLTREAGSTVMWAGQEVYVICVS</td> <td>-----</td> <td>CGAKP-----APDII</td> <td>172</td>	119	GPELVSRVILSVLVPRKVLQLTREAGSTVMWAGQEVYVICVS	-----	CGAKP-----APDII	172
QY <td>162</td> <td>NIQGGKTL</td> <td>-----</td> <td>SVQLELQDSGTWTCTVYLQNOK</td> <td>195</td>	162	NIQGGKTL	-----	SVQLELQDSGTWTCTVYLQNOK	195
DB <td>173</td> <td>FICGGRTVEDSSVNEGSEBEKLFTEAEARVTPQSSDNQGLLVCESNPALATPIKASF</td> <td>-----</td> <td></td> <td>232</td>	173	FICGGRTVEDSSVNEGSEBEKLFTEAEARVTPQSSDNQGLLVCESNPALATPIKASF	-----		232
QY <td>196</td> <td>KIDIVPCAPRPSKCDKHTHCPELLGGSVLFPPKPRDITLMISRTPEVTCVVDVSHED</td> <td>-----</td> <td></td> <td>255</td>	196	KIDIVPCAPRPSKCDKHTHCPELLGGSVLFPPKPRDITLMISRTPEVTCVVDVSHED	-----		255
DB <td>233</td> <td>TNLTLP</td> <td>-----</td> <td>GPVTDWGLNEGHVRAGENELPCT--ARGN</td> <td>271</td>	233	TNLTLP	-----	GPVTDWGLNEGHVRAGENELPCT--ARGN	271
QY <td>256</td> <td>PEVEFNWVDSGEVHNAAKTKPREQVNSYRVSVLFTVLDHDMNGEKYKCKVSNAKALPA</td> <td>-----</td> <td></td> <td>315</td>	256	PEVEFNWVDSGEVHNAAKTKPREQVNSYRVSVLFTVLDHDMNGEKYKCKVSNAKALPA	-----		315
DB <td>272</td> <td>PPALQLHKNKPKPSIAW</td> <td>-----</td> <td>GTEHAQAVNASVLVMTYRPED--HGARLSQCSYNSVAE</td> <td>326</td>	272	PPALQLHKNKPKPSIAW	-----	GTEHAQAVNASVLVMTYRPED--HGARLSQCSYNSVAE	326
QY <td>316</td> <td>PIEKTISKAKQAPREPOVYTLPPSRDELITKNQVSLTCLVKGFPYSIDIAVEWESNGQPENN</td> <td>-----</td> <td></td> <td>375</td>	316	PIEKTISKAKQAPREPOVYTLPPSRDELITKNQVSLTCLVKGFPYSIDIAVEWESNGQPENN	-----		375
DB <td>327</td> <td>TOERSITLQVTFP--PSAVITLIGSTQSEKNQVTLCCITKSRRP-VLLRWMLGR--QL</td> <td>-----</td> <td></td> <td>381</td>	327	TOERSITLQVTFP--PSAVITLIGSTQSEKNQVTLCCITKSRRP-VLLRWMLGR--QL	-----		381
QY <td>376</td> <td>YKTPPVLDS-DGSEFLYSKLTVDKSRQGNVNSCSVMHEALNNHTYQKSLSPGLQL</td> <td>-----</td> <td></td> <td>434</td>	376	YKTPPVLDS-DGSEFLYSKLTVDKSRQGNVNSCSVMHEALNNHTYQKSLSPGLQL	-----		434
DB <td>382</td> <td>LPTBETVWDGHLGHGISMSNTLLTVKREDNGLSLTCEAFSAFKETFPKSLTLNVKXPA</td> <td>-----</td> <td></td> <td>441</td>	382	LPTBETVWDGHLGHGISMSNTLLTVKREDNGLSLTCEAFSAFKETFPKSLTLNVKXPA	-----		441
QY <td>435</td> <td>DETCAEAQDGLDGLMTTDPRAASL</td> <td>-----</td> <td></td> <td>460</td>	435	DETCAEAQDGLDGLMTTDPRAASL	-----		460
DB <td>442</td> <td>OK-----LWIBGPEGQSI</td> <td>-----</td> <td></td> <td>455</td>	442	OK-----LWIBGPEGQSI	-----		455

RESULT 93  
 LAC1 RAT STANDARD; PRT: 104 AA.

AC P20766;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ig lambda-d1 chain C region.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NC Mammalia; Eutheria; Rodentia; Sciurgnathi; Muridae; Murinae; Rattus.  
 NCBI taxid=10116;





```

FT CARBOHYD 1029 1029 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1106 1106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1179 1182 Missing (in isoform 2).
FT SEQUENCE 1259 AA; 140934 MW; 0F12A7C4415F3C08 CRC64;

Query Match
Best Local Similarity 24.4%; Pred. No. 0.1;
Matches 99; Conservative 43; Mismatches 138; Indels 126; Gaps 24;

QY 10 LLLVQLALLPAATGKNVVLGKGDVETLTASQKSIQ--FWKNSNQIKILGNQGS 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 507 ILANLOVKATQITQGPSTIEKKGARVFTQASDPSTLQASITWRDGR----- 557
QY 68 FLTKGSKLNDPRDSRLMDQGNFLLIKNLKIEBSPYIC----EVDQKEEVQLLVF 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 558 -----DLQERKDSKRYFIEDQ--LVIKSLDYSQDGYSCVASTELDEVSRALLV 608
QY 124 GLTAN-----SPTHLLQGSLTLTLESPGSSPSVQCRRSP-----RGKNIQGGKTL 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 609 GSRGPRPHLELSRHLIKOSQVHLN-----SPADHNSPIEKYDIEFEDKEMAEKMF 662
QY 170 SVSQLELDQSGTWCTVLQNKVVE--FKIDIVPCBP-EPKSCDTHTCPELLGSPVF 226
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 663 SLGKV-----PQNGTSTTLKSLPYVHYTFRTAIKNGPGEPSVSTVTPPE----- 710
QY 227 LPPPKKDTLMISRTPEVTCVVYDVSHDEPEVK--FNNYVDGCVNAKATRP-REEQYN 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 711 -----AAPEKN--PVDVREGNETNMVITW-----KPLRMMDNN 743
QY 283 S--TYRVAVSLTVLHQMNLN-GKEYKCK-----VSNKALPAPIEKTISKAKGQP 328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 744 AAOIQIRV-----QMRPLCKQETWKEQVYSDPFLVVSNTSTVPVLEIKYQAVNNQG 794
QY 329 R--EPQV-----YTP-----LPPSRDELTKNQVSLTCLVKGFPSDIA 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 795 KQREPGVTIGSGEDYPOVSPLEEDITTFN-SSTVLVR-WRPVDLA 838

RESULT 95
LAC_CHICK
ID LAC_CHICK STANDARD; PRT; 103 AA.
AC P20763;
DT 01-FEB-1991 (rel. 17, Created)
DT 01-FEB-1991 (rel. 17, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Ig lambda chain C region.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN (1)
RX MEDLINE=87218480; PubMed=3107981;
RA Parvart R., Ziv E., Lenter F., Tel-Or S., Bursstein Y., Schechter I.;
RT "Analyses of chicken immunoglobulin light chain cDNA clones indicate
a few germline V lambda genes and allotypes of the C lambda locus.";
RL EMBO J. 6:97-102(1987).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC EMBL; X04768; CAA28461.1; -
CC PIR; B26167; B26167.
CC HSSP; P01842; 7FAB.

```

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00407; IgC1; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 6 99 IG-LIKE.
FT DISULFID 28 85
FT DISULFID 103 103 INTERCHAIN (WITH HEAVY CHAIN).
FT VARIANT 90 90 N -> D.
SQ SEQUENCE 103 AA; 11361 MW; 77BF341B511B91B2 CRC64;

Query Match
Best Local Similarity 32.2%; Pred. No. 0.0048;
Matches 29; Conservative 19; Mismatches 39; Indels 3; Gaps 3;

QY 327 QPR-EPQVYTLPPSRDELTK-NQVSLTCLVKGFPSPDIAVEMESNGQENNYKTPPYLD 384
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 QPKVAPITLTPPSKELEATKATLVCLINDFPSPVTVDWIDGSTRSG-ETTAPOKQ 59
QY 385 SDGSFFLYSKLTVDSRMQGNVPSQYWH 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 60 SNSQYMASSTYLSASDWSHETTCRYTH 89

RESULT 96
KACB_RAT
ID KACB_RAT STANDARD; PRT; 106 AA.
AC P01835;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, Last sequence update)
DT 10-OCT-2003 (rel. 42, Last annotation update)
DE Ig kappa chain C region, B allele.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RX MEDLINE=82082587; PubMed=6273908;
RA Sheppard H.W., Gutman G.A.;
RT "Allelic forms of rat kappa chain genes: evidence for strong
selection at the level of nucleotide sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:7064-7068(1981).
CC [2]
CC SEQUENCE (BENCE-JONES PROTEIN S211).
RX MEDLINE=75212238; PubMed=807630;
RA Starace V., Querinjean P.;
RT "The primary structure of a rat kappa Bence Jones protein:
phylogenetic relationships of V- and C-region genes.";
RL J. Immunol. 115:59-62(1975).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC PIR; A93901; KIRTB.
DR HSSP; P01842; 2MCG.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00407; IgC1; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 5 102 IG-LIKE.
FT DISULFID 26 86
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT CONFLICT 2 2 D -> N (IN REF. 2).
FT CONFLICT 30 30 N -> K (IN REF. 2).
FT CONFLICT 48 48 MISSING (IN REF. 2).

```

```

FT CONFLICT 79 79 E -> Q (IN REF. 2).
FT CONFLICT 87 87 E -> Q (IN REF. 2).
FT CONFLICT 98 98 V -> VM (IN REF. 2).
FT CONFLICT 100 100 S -> N (IN REF. 2).
SQ SEQUENCE 106 AA: 11601 MW: 4CFA7CA820D1CA36 CRC64:

Query Match
Best Local Similarity 31.5%; Pred. No. 0.0053;
Matches 29; Conservative 22; Mismatches 29; Indels 12; Gaps 2;

QY 331 PQTTLPSRDELTONQVSLCLVKGFPSPDINAVEMENSCPENNYKTPVLD-----385
DB 5 PTVSIFPSTQLATGASVCLMKNFYPRDISVYKWKIDGERD-----GVLSDVTDQD 59
QY 386 --DGSFPLYSKLTVDKSRMOQGVNFCVME 415
DB 60 SKDSTYSMSSTLSLTAKDYESHNLTYTCVYVK 91

RESULT 97
AMAL_DROME STANDARD; PRT: 333 AA.
ID AMAL_DROME
AC P15364; Q9V3A5;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Amalgam protein precursor.
GN AMA OR BG:DS00276.6 OR CG2198.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=89028670; PubMed=3141062;
RA Seeger M.A., Haffley L., Kaufman T.C.;
RT "Characterization of amalgam: a member of the immunoglobulin
RT superfamily from Drosophila.";
RL Cell 55:589-600(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Celniker S.E., Pfeiffer B.D., Knafela J., Martin C.H., Mayeda C.A.,
RA Palazolo M.J.;
RT "Complete sequence of the Antennapedia complex of Drosophila.";
RT submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.W., Baau P.V., Berman B.P., Bhandari D., Bolshakov S.M.,
RA Beeson K.Y., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Butts J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fowler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Idbegam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemtson J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobary C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Puri V., Reese M.G.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasmann D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guerin H., Krommiller B., Pacheb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (potential).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M23561; AA28367.1; -.
DR EMBL: AE001572; AAD19797.1; -.
DR EMBL: AE003674; AAF54084.1; -.
DR EMBL: AY051911; AAK93335.1; -.
DR PIR: A31923; A31923.
DR FLYBase: FBgn0000071; Ama.
DR GO: GO:0005886; C:Plasma membrane; IDA.
DR InterPro: IPR007110; IG_1like.
DR InterPro: IPR003598; IG_C2.
DR Pfam: PF00047; IG_3.
DR SMART: SM00408; IGC2; 2.
DR PROSITE: PSS0835; IG_LIKE; 3.
KW Immunoglobulin domain; Glycoprotein; Membrane; GPI-anchor; Signal;
KW Repeat.
FT 1 23 POTENTIAL.
FT CHAIN 24 ? AMALGAM PROTEIN.
FT PROPEP 2 333 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 25 128 IG-LIKE V-TYPE.
FT DOMAIN 139 223 IG-LIKE C2-TYPE 1.
FT DOMAIN 230 323 IG-LIKE C2-TYPE 2.
FT DISULFID 46 117 PROBABLE.
FT DISULFID 161 208 PROBABLE.
FT DISULFID 251 307 PROBABLE.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 83 83 Q -> K (IN REF. 1).
SQ SEQUENCE 333 AA: 36387 MW: F644753DE3DB25F1 CRC64;

Query Match
Best Local Similarity 19.9%; Pred. No. 0.024;
Matches 70; Conservative 57; Mismatches 127; Indels 97; Gaps 14;

```

QY 10 LLLVLQAL-----LPAATQGNKRVVLGKKGDVETLCTASQKSIQPHW-----KNSNQ 58

```

Db      10 LIFCLALSLDLSVLSAPVLSQISKDVASVSDSVFNPCTVEVQGLSVNAKRSRSDTNS 69
Qy      59 IKI-----LGNQ--GSFTRKPSKLNDRADSRSLMDQGNPLLIKLNKIDSDTYIC 109
Db      70 VILSMENISLSPQORVNVVTEGPKT-----GSAITFRIONIESVDMQMPYEC 117
Qy      110 EV-----EDQKEVQL--VFGLTNSDTHLLOQOSLTLTLESPGSSPSV----- 153
Db      118 QVLVSATKTKLKLQIKPPVIAENTPKSTLVTEGQNLCHANGPKPTISWAREH 177
Qy      154 QCRSPGKNIQGGKTLVSQLELQDSGTWTCTVLON-----QKVEFKIDIVPC 203
Db      178 NAWMPAGHLLAPPTLRIRSVHMDGYYC-IAQNGEQPDRLRVEVEFRPQLA--- 233
Qy      204 APEPKGCDKTHTEPELIGGSVLFPPPKQDTMISTRTPTVTVVVDVSHEDPVKFNMY 263
Db      234 -----VORPKIAGMVSLSLECSVOGY-----PAPTVVWH 264
Qy      264 VDGVEVNAKTKPREQVNSTVRVSVL---TVLHDMVNGKXKCKVSNX 311
Db      265 KNGVPLDSSRHHEVANTASSGTTTSVLRLDSVGEEDF---GDYCNATWK 312

RESULT 98
VCAM_MOUSE
ID VCAM_MOUSE STANDARD: PRT: 739 AA.
AC P29533;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular cell adhesion protein 1 precursor (V-CAM 1).
GN VCAM1 OR VCAM-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid:10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=FVB; TISSUE=Lung;
RX MEDLINE=92161437; Pubmed=1371918;
RA Hession C., Moy P., Tizard R., Chieholm P., Williams C., Wyck M.,
RA Burkly L., Miyake K., Kincaid P., Lobb R.;
RT "Cloning of murine and rat vascular cell adhesion molecule-1.";
RL Biochem. Biophys. Res. Commun. 183:163-169(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lymph node;
RX MEDLINE=93246254; Pubmed=7683304;
RA Arai M., Arai K., Vassalli P.;
RT "Cloning and sequencing of mouse VCAM-1 cDNA.";
RL Gene 126:261-264(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=129; TISSUE=Embryo;
RA Ballantyne C.M.;
RT Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-693 FROM N.A. (ISOFORM 1).
RC STRAIN=NIH Swiss, and 129/Sv;
RA Kumar A.G., Dai Y.X., Kozak C.A., Mims M.P., Gotto A.M. Jr.,
RA Ballantyne C.M.;
RT Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=FVB; TISSUE=Lung;
RX MEDLINE=93232042; Pubmed=7682556;
RA Moy P., Lobb R., Tizard R., Olson D., Hession C.;
RT "Cloning of an inflammation-specific phosphatidylinositol-linked
RT form of murine vascular cell adhesion molecule-1.";
RL J. Biol. Chem. 268:8835-8841(1993).

```

```

RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=95015899; Pubmed=7523515;
RA Kumar A.G., Dai Y.X., Kozak C.A., Mims M.P., Gotto A.M.,
RA Ballantyne C.M.;
RT "Murine VCAM-1. Molecular cloning, mapping, and analysis of a
RT truncated form.";
RL J. Immunol. 153:4088-4098(1994).
RN [7]
RP SEQUENCE OF 311-345 FROM N.A. (ISOFORM 2).
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=93317595; Pubmed=7687058;
RA Terry R.W., Kwee L., Levine J.F., Labow M.A.;
RT "Cyclokin induction of an alternatively spliced murine vascular cell
RT adhesion molecule (VCAM) mRNA encoding a
RT glycosylphosphatidylinositol-anchored VCAM protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5919-5923(1993).
RN [8]
RP SEQUENCE OF 1-21 FROM N.A.
RC TISSUE=Endothelial cells;
RA Korenaga R., Ando J., Teubol H., Kamiya A.;
RT Submitted (Dec-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
CC IN LEUCOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
CC INTEGRIN VIA4 ON LEUCOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
CC TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A
CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUCOCYTE
CC EMIGRATION TO SITES OF INFLAMMATION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG ISOFORM) OR
CC ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (SHORT ISOFORM).
CC -1- ALTERNATIVE PRODUCTS:
CC Name=1; Synonyms=Long;
CC Name=2; Synonyms=Short;
CC IsoId=P29533-1; Sequence=Displayed;
CC IsoId=P29533-2; Sequence=VSP 002581, VSP 002582;
CC Name=2; Synonyms=Short;
CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as
CC well as on macrophage-like and dendritic cell types in both normal
CC and inflamed tissue.
CC -1- PTM: The GPI-anchor is located on position 319 of isoform 2.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M8487; AAA40545.1; -
CC EMBL; X67783; CAA47989.1; -
CC EMBL; L22355; AAA16921.1; -
CC EMBL; L22301; AAA16921.1; JOINED.
CC EMBL; L22349; AAA16921.1; JOINED.
CC EMBL; L22350; AAA16921.1; JOINED.
CC EMBL; L22351; AAA16921.1; JOINED.
CC EMBL; L22352; AAA16921.1; JOINED.
CC EMBL; L22353; AAA16921.1; JOINED.
CC EMBL; L22354; AAA16921.1; JOINED.
CC EMBL; L22355; AAA16921.1; JOINED.
CC EMBL; L22350; AAA16920.1; -
CC EMBL; L22301; AAA16920.1; JOINED.
CC EMBL; L22349; AAA16920.1; JOINED.
CC EMBL; L22879; AAB60659.1; ALT_SEQ.
CC EMBL; L22879; AAB60660.1; ALT_SEQ.
CC EMBL; L22880; AAB60661.1; ALT_SEQ.
CC EMBL; L22874; AAB60662.1; ALT_SEQ.
CC EMBL; L22871; AAB60663.1; ALT_SEQ.
CC EMBL; L22883; AAB60664.1; ALT_SEQ.
CC EMBL; L22881; AAA80010.1; ALT_SEQ.
CC EMBL; L22882; AAA80011.1; ALT_SEQ.
CC EMBL; L22875; AAA80012.1; ALT_SEQ.

```

[illegible]

```

Oy 248 VVDV---SHEPPE-----VKFMWYDGVVHNAKTPREEOY-----NSTRVVSVLTV 299
Db 427 VPNVYPPDHEIELLKGETTLMKRYFLEPMQIKSLFTLTFTPTIEDTKSLVCLAR 486
Oy 294 LHODLMNGKEYCKVSKNALPAPIEKTIKSAKGOPREPOVYTLR-PSRDELTKNOVSLTC 352
Db 487 LHSG-----EMESBPQRQSVQPLVYVNA-----PKETIIWVSPSILIE--GSPVNLTC 534
Oy 353 LVKGFYSYDLAVE---SNQPENNYKTPPVLDSDSFLYKSLTYDKSKWQGNVFS 409
Db 535 SSDGIPIAPKIL--WSROLNNGELQ-----PLESN-----TTLTPTMSTRKDSGIYV 578
Oy 410 CSVMHEALHNHYTQKSL 428
Db 579 CEGINEA---GISRKSVEL 594

RESULT 99
LAMP RAT ID LAMP RAT STANDARD; PRT; 338 AA.
AC Q62813;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP) .
GN LSAMP OR LAMP.
OS Rattus norvegicus (Rat) .
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116,
[1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-49.
RC TISSUE=Hippocampus;
RX MEDLINE=95374785; PubMed=7646886;
RA Pimenta A.F., Zhukareva V., Barbe M.F., Reinoso B.S., Grimley C.,
RA Henzel W., Fischer I., Levitt P.;
RT "The limbic system-associated membrane protein is an Ig superfamily
member that mediates selective neuronal growth and axon targeting.";
RL Neuron 15:287-297(1995) .
CC -1- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
CC NATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: Expressed mostly by neurons comprising limbic-
CC associated cortical and subcortical regions that function in
CC cognition, emotion, memory, and learning.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT E15-16, AT STAGE E20 IT IS
CC DETECTED IN PRESUMPTIVE CORTEX, MEDIAL LIMBIC AREAS OF THE
CC THALAMUS AND HYPOTHALAMUS. IN THE ADULT, IT IS FOUND IN
CC HYPOTHALAMUS, PERIRHINAL CORTEX, AMYGDALA AND MEDIAL THALAMIC
CC REGION.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce
CC or send an email to license@isb-sib.ch ).
CC
CC EMBL: U31554; AAA86120.1; -.
CC InterPro: IPR007110; Ig-like.
CC InterPro: IPR003598; Ig_c2.
CC Pfam: PF00047; Ig; 3.
CC SMART: SM00408; IGC2; 2.
CC SMART: PS50835; IG LIKE; 3.
CC Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
CC Repeat; Signal; Lipoprotein.
CC SIGNAL
CC 1
CC 28

```

FT	CHAIN	29	315	LIMBIC SYSTEM-ASSOCIATED MEMBRANE PROTEIN REMOVED IN MATURE FORM (POTENTIAL).
FT	PROPEP	316	338	IG-LIKE C2-TYPE 1.
FT	DOMAIN	29	122	IG-LIKE C2-TYPE 1.
FT	DOMAIN	132	214	IG-LIKE C2-TYPE 2.
FT	DOMAIN	219	304	IG-LIKE C2-TYPE 3.
FT	DISULFID	53	111	POTENTIAL.
FT	DISULFID	153	197	POTENTIAL.
FT	DISULFID	239	290	POTENTIAL.
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	66	66	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	148	148	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	279	279	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	287	287	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	300	300	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	315	315	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	LIPID	315	315	GPI-anchor amidated asparagine (Potential).
SO	SEQUENCE	338 AA;	37324 MW;	0B76AFDD68A39B86 CRC64;
	Query Match	5.3%;	Score 144;	DB 1; Length 338;
	Best Local Similarity	22.4%;	Pred. No. 0.026;	
	Matches	74;	Conservative	50; Mismatches 129; Indels 78; Gaps 16.
QY	10	LLVLVLQLALLPA-----TQGNKVVVLGKKGKDTVELLTCTASQKKSIOFHKNSNQIKI	61	
DB	14	LVLRLRLCLTPTGLFPRSVDFENRGDNTITVQGDPAILRCVVEDKNS-KVAMLNRSGLIF	72	
QY	62	LGNQGSFLITGKPSKLNDR--DSRSINLQDGNFPIILNKIKEDSDTYICEVEDKE---	116	
DB	73	AGHD-----KMSLDIPRVELEKRRAL---EYSLRIRKVDYDGGSYTCSVOTQHEKPT	121	
QY	117	-EVOLLVG-----LTPNSDTHLLOQSITLTLSPGSSPSVQCR--SPRGNIQGKT-	168	
DB	122	SOVLIVGVPRKISNISSDVYNBSNTVLICMANGREPVITWHLTPLRFPFGEEBY	181	
QY	169	LSVSQLELDQSGTWTCTVLQ-----NOKKVEKIDIVCPAPEKSCDKTHTCPELLGAP	223	
DB	182	LEIGITREGSGKCEKCAANEVSADVKQVKTAVNYP-TITESKSNAT-----	230	
QY	224	SVFLFPPPKPKDTLMISRTPEYTCVVVDVSHEDPEVKFMWYVDGVGVAHA--KTYPREQ	280	
DB	231	-----TGROASLKCEASAV-----PADDFEWMYRDTDRINSANGLEIKSTEGQ	272	
QY	281	YNSTRVYVSLTVLHODWLNGKEYCKYSNK	311	
DB	273	SSLT-----VTVNTEHYGN---YTCAVANK	295	
	RESULT 100			
	NEFAS_RAT	STANDARD:	PRT:	1240 AA.
AC	P97685; P97684; Q91260;			
DT	15-MAR-2004 (Rel. 43, Created)			
DT	15-MAR-2004 (Rel. 43, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Neurofascin precursor.			
GN	NPASC.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	SEQUENCE FROM N.A. (ISOFORM 3).			
RP	STRAIN=H1starc;			
RC	Tat1 S., Collinson J.M., Brophy P.J.;			
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.			
RN	SEQUENCE OF 25-1240 FROM N.A. (ISOFORMS 1; 2 AND 3), AND SUBUNIT.			
RC	Tissue=Brain;			
EX	MEDLINE=97103184; PubMed=8947556;			
RA	Davis J.O., Lambert S., Bennett V.;			

RT "Molecular composition of the node of Ranvier: identification of  
RT ankyrin-binding cell adhesion molecules neurofascin (mucin+/chird  
RT FNIII domain-) and NrCAM at nodal axon segments.";  
RL J. Cell Biol. 135:1355-1367(1996).  
RN [3]  
RP FUNCTION OF ISOFORM 2/3.  
RA MEDLINE=98220650; Pubmed=9562181.  
RX Collinson J.M., Marshall D., Gillespie C.S., Brophy P.J.;  
RT "Transient expression of neurofascin by oligodendrocytes at the onset  
RT of myelogenesis: implications for mechanisms of axon-glia  
RT interaction.";  
RL J. Neurosci. 17:1123-1131(1997).  
RN [4]  
RP FUNCTION OF ISOFORMS 1 AND 2/3, INDIRECT ASSOCIATION WITH CNTNAP1,  
RP PHOSPHORYLATION, AND TISSUE SPECIFICITY.  
RX MEDLINE=20391985; Pubmed=10931875;  
RA Tait S., Gunn-Moore F., Collinson J.M., Huang J., Lubetski C.,  
RA Pedraza L., Sherman D.L., Colman D.R., Brophy P.J.;  
RT "An oligodendrocyte cell adhesion molecule at the site of assembly of  
RT the paranodal axo-glia junction.";  
RL J. Cell Biol. 150:657-666(2000).  
RN [5]  
RP INTERACTION WITH SODIUM CHANNEL BETA-1 AND BETA-3 SUBUNIT.  
RX MEDLINE=21365577; Pubmed=11470823;  
RA Raccaliffe C.F., Westendorp R.E., Curtis R., Caterall W.A.;  
RT "Sodium channel beta1 and beta3 subunits associate with neurofascin  
RT through their extracellular immunoglobulin-like domain.";  
RL J. Cell Biol. 154:427-434(2001).  
CC -1- FUNCTION: Cell adhesion, ankyrin-binding protein which may be  
CC involved in neurite extension, axonal guidance, synaptogenesis,  
CC myelination and neuron-glia cell interactions. Isoforms 2/3 may  
CC be responsible for mediating and signaling axon-glia interaction  
CC during the early stages of myelination.  
CC -1- SUBUNIT: Probable constituent of a neurofascin/NRCAM/ankyrin G  
CC complex. Associates with the sodium channel beta-1 (SCN1B) and  
CC beta-3 (SCN1B) subunits. Associates to beta-1 subunit in  
CC developing axons as early as postnatal day 5, during the period  
CC that nodes of Ranvier are forming. Isoform 2/3 is likely to  
CC interact with axonal proteins in close association with CNTNAP1.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 1  
CC colocalizes with ankyrin G at the nodes of Ranvier. Isoform 2/3 is  
CC a glial component of the paranodal axo-glia junction.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named Isoforms=3;  
CC Name=1; Synonyms=NF186, 186 kDa isoform;  
CC IsoId=P97685-1; Sequence=D13displayed;  
CC Name=2; Synonyms=NF155, 155 kDa isoform;  
CC IsoId=P97685-2; Sequence=VSP\_050416, VSP\_050417, VSP\_050418,  
CC VSP\_050419;  
CC Name=3; Synonyms=NF155, 155 kDa isoform;  
CC IsoId=P97685-3; Sequence=VSP\_050416, VSP\_050417, VSP\_008941,  
CC VSP\_050418, VSP\_050419;  
CC -1- TISSUE SPECIFICITY: Isoform 1 is expressed at Nodes of Ranvier  
CC while isoform 2/3 is expressed in unmyelinated axons.  
CC -1- DEVELOPMENTAL STAGE: Strongly but transiently up-regulated in  
CC oligodendrocytes at the onset of myelogenesis. Once these last  
CC have engaged their target axons, expression declines  
CC precipitously.  
CC -1- PTM: Isoform 2/3 is phosphorylated at P12. Dephosphorylation is  
CC required for ankyrin binding.  
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.  
CC LI/neurofascin/NGCAM family.  
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.  
CC -1- SIMILARITY: Contains 4 fibronectin type III domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AY061639; AAL27854.1; -.  
DR EMBL; U81035; AAB47753.1; -.  
DR EMBL; U81036; AAB47754.1; -.  
DR HSPSP; P20241; 1CER.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR008957; FN\_III-like.  
DR InterPro; IPR007110; IG\_1like.  
DR InterPro; IPR003598; IG\_C2.  
DR Pfam; PR00047; fn3; 4.  
DR Pfam; PR00047; fn3; 6.  
DR SMART; SM00060; FN3; 3.  
DR SMART; SM00408; IGC2; 4.  
DR PROSITE; PS00835; IG\_LIKE; 6.  
DR Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;  
KW Glycoprotein; Alternative splicing; Phosphorylation.  
FT SIGNAL 1 24  
FT CHAIN 25 1240  
FT DOMAIN 25 1110  
FT TRANSMEM 1111 1131  
FT DOMAIN 1132 1240  
FT DOMAIN 41 137  
FT DOMAIN 143 230  
FT DOMAIN 244 332  
FT DOMAIN 337 424  
FT DOMAIN 429 517  
FT DOMAIN 521 603  
FT DOMAIN 628 720  
FT DOMAIN 727 820  
FT DOMAIN 825 918  
FT DOMAIN 1008 1094  
FT DOMAIN 913 1006  
FT DISULFID 63 118  
FT DISULFID 162 213  
FT DISULFID 268 316  
FT DISULFID 358 408  
FT DISULFID 452 501  
FT DISULFID 543 592  
FT CARBOHYD 305 305  
FT CARBOHYD 409 409  
FT CARBOHYD 446 446  
FT CARBOHYD 483 483  
FT CARBOHYD 752 752  
FT CARBOHYD 778 778  
FT CARBOHYD 866 866  
FT CARBOHYD 881 881  
FT CARBOHYD 1015 1015  
FT CARBOHYD 1026 1026  
FT CARBOHYD 1047 1047  
FT CARBOHYD 1101 1101  
FT VARSPLIC 31 36  
FT VARSPLIC 236 236  
FT VARSPLIC 611 625  
FT VARSPLIC 824 824  
FT VARSPLIC 928 1096  
FT CONFLICT 482 482  
FT CONFLICT 675 675  
FT CONFLICT 699 699  
FT CONFLICT 763 763  
SQ SEQUENCE 1240 AA; 138003 MW; 636A187BC3772513 CRC64;  
Query Match 5.3%; Score 143.5; DB 1; Length 1240;  
Best Local Similarity 21.4%; Pred. No. 0.15;  
Matches 106; Conservative 67; Mismatches 193; Indels 129; Gaps 28;

QY 6 PFRHLVLVQLALLPAA-----TQGNKVVIGKKKGDVVELTCTASQ 45  
DB 8 PFWVALILFLSLGALIEIPMPSIONELTQPTITKQSVKDIIVDRNIIILECEAKG 67  
QY 46 KKSIOFHW-KKSNQIKILGNQGSFLTK--GSKLNDPADSRSLMDQGNFLLIKN----- 98  
DB 68 NPASFHWTRNSRFNIAKQPRVSMRRSGTLVIDFSGGRPEBY-EGEYOCFARNKFGT 126  
QY 99 -----LKIEDSDTYICEVEDQEEVQLVFGLTANSDTLLQGGSTLTLLESPPG-SS 150  
DB 127 ALSNRIRLQVSKSLW-----PKENLDPV-----VQSGAPLTLQCNPPGLPS 170  
QY 151 PSVQCRSPRGKNIQGGKTLG-----VSQLELDQSGT-WTCTVLQN-----QKVEFK 196  
DB 171 FVIFMSSMSEPIIQDKRVSGHNGDLVFSNVMQLDMQTDVSCNARFHTHTIQKXNFT 230  
QY 197 IDIYPCAPPEKSCDKHTCTCELLGSGVFLFPKPKDITLMISTPE--VTCVVVDVSH 254  
DB 231 LKVLITRGVAERT-----PS-FMYPGTSSSQVLRGMDDLLECLASGV--- 273  
QY 255 DPEYKFWYVDGVFNHNAKTPREEOYNSYTRVSVLTVLHQMVLNGKEYCKVSNKALP 314  
DB 274 -PTVDIANYKKKGDLPSDKAK--FENFNKALRITNVSEB-----DSGEYFCLASNGM-- 322  
QY 315 APIEKTIS-KAKGP--REPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVESWNG 370  
DB 323 GSIRHTISVRKAPYWLDEPKNLIAPGEDG-----RLVCBRANG-NPKP-TVQWLVNG 374  
QY 371 QPENNYKTPPVLDSD--GSFPLYSKLTVDKSRMOQGVFSCSVNHEALNHHYQKSL-- 426  
DB 375 DP--LQSAFPNPFEVAGDTIIFRDQISSRA-----VYQCNISNE--HGVLNANAFVS 424  
QY 427 -----SLSPGLQ 434  
DB 425 VLDPVPRMLSPRNQ 439

Search completed: August 3, 2004, 13:09:30  
Job time : 13.0314 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 13:00:24 ; Search time 11.7668 Seconds  
(without alignments)  
4185.504 Million cell updates/sec

Title: SE08  
Perfect score: 2702  
Sequence: 1 NMRGVPRHLLVLQLALP.....VISFLGLGVACVLAARTR 512

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 125 summaries

Database :  
1: pir78:\*  
2: pir1:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1227.5	45.4	330	1 GHU	Ig gamma-1 chain C
2	1221	45.2	374	2 S69339	Ig heavy chain V r
3	1203.5	44.5	255	4 S31866	Ig gamma-1 chain C
4	1173.5	43.4	234	2 PTO207	Ig gamma chain C r
5	1157	42.8	377	2 A23511	Ig gamma-3 chain C
6	1155	42.7	377	2 A60764	Ig gamma-3 chain C
7	1133	41.9	289	1 G3HOW1	Ig gamma-3 heavy C
8	1130.5	41.8	326	1 G2HU	Ig gamma-2 chain C
9	1119	41.4	327	1 G4HU	Ig gamma-4 chain C
10	1023	37.9	458	1 RMHT4	T-cell surface gly
11	955.5	35.4	398	1 G3MSM	Ig gamma-3 chain C
12	932.5	34.5	393	1 G3MSM	Ig gamma-1 chain C
13	913.5	33.9	470	2 S22080	Ig heavy chain pre
14	906.5	33.5	399	1 G2MSAM	Ig gamma-2a chain
15	901	33.3	328	2 I47159	Ig gamma-2a chain
16	899	33.3	323	1 GHRB	Ig gamma chain C r
17	898	33.2	327	2 I47162	Ig gamma 4 chain c
18	898	33.2	328	2 I47160	Ig gamma 2b chain
19	894.5	33.1	329	1 G2GP	Ig gamma-2 chain C
20	879.5	32.5	405	1 I47158	Ig gamma-1 chain C
21	877.5	32.4	432	1 G2MSBM	Ig gamma-2b chain
22	875	32.4	432	1 RMCZT4	T-cell surface gly
23	873	32.3	328	2 I47161	Ig gamma 3 chain c
24	857.5	31.7	329	1 G3MSC	Ig gamma-3 chain C
25	854.5	31.6	444	2 PC4436	monoclonal antibod
26	854.5	31.6	472	2 S31459	Ig gamma-1 chain -
27	850.5	31.5	469	2 S37483	Ig gamma-2a chain
28	838	31.0	308	2 C30554	Ig heavy chain C r
29	833	30.8	333	2 PS0018	Ig gamma-2b chain

30	825	30.5	446	2 S40295	Ig gamma-2a chain
31	823.5	30.5	326	2 PS0017	Ig gamma-1 chain C
32	823	30.5	329	2 S00847	Ig gamma-2c chain
33	818.5	30.3	324	1 G3MS1	Ig gamma-1 chain C
34	818	30.3	474	1 G2MS11	Ig gamma-2b chain
35	805.5	29.8	330	1 G2MSA	Ig gamma-2a chain
36	801	29.6	335	1 G2MSAB	Ig gamma-2a chain
37	793.5	29.4	475	2 S01321	Ig gamma-2b chain
38	792.5	29.3	322	2 PS0019	Ig gamma-2a chain
39	790	29.2	432	1 RMWQ14	T-cell surface gly
40	774	28.6	327	2 S06611	Ig gamma-2 chain C
41	702	26.0	180	2 I46732	Ig gamma heavy cha
42	601.5	22.3	459	2 A46254	CD4 precursor - ra
43	581.5	21.5	432	2 S30193	T-cell surface gly
44	574.5	21.3	218	2 A36040	Ig heavy chain V-I
45	572.5	21.2	249	2 S69340	Ig heavy chain VHI
46	566	20.9	152	2 S14236	Ig gamma-1 chain C
47	495	18.3	457	2 A27449	T-cell surface gly
48	475	17.6	572	2 B46529	T-cell surface gly
49	411.5	15.2	592	2 S04845	Ig Y heavy chain (
50	402.5	14.9	549	2 S04845	Ig heavy chain pre
51	389	14.4	627	2 S14683	Ig mu chain precu
52	388	14.4	388	1 EHMS	Ig epsilon chain C
53	386.5	14.3	474	2 S15590	Ig heavy chain - h
54	385	14.2	453	2 S37768	Ig mu chain C regi
55	384	14.2	592	2 S25705	Ig mu chain - shee
56	381	14.1	473	1 MHRDM	Ig mu chain C regi
57	381	14.1	548	2 S38864	Ig epsilon chain C
58	376.5	13.9	452	1 MHRU	Ig mu chain C regi
59	374	13.8	448	2 S03186	Ig heavy chain C regi
60	373.5	13.8	476	1 MHRMS	Ig mu chain C regi
61	370.5	13.7	429	1 EHRT	Ig epsilon chain C
62	369.5	13.7	479	1 MHRBM	Ig mu chain C regi
63	369	13.7	450	1 MHDG	Ig mu chain C regi
64	365.5	13.5	391	1 MHRUBT	Ig mu heavy chain
65	363.5	13.4	423	1 EHMS	Ig epsilon chain C
66	363	13.4	455	2 A24976	Ig mu chain C regi
67	362	13.4	455	1 MHRMS	Ig mu chain C regi
68	361.5	13.4	426	2 I36948	Ig epsilon chain -
69	361.5	13.4	428	1 EHRU	Ig epsilon chain C
70	361	13.4	504	2 S00390	Ig gamma chain (c1
71	360	13.3	458	1 MHRB	Ig mu chain C regi
72	357	13.2	71	2 I60082	CD4 receptor - hum
73	354.5	13.1	457	2 S03961	Ig mu chain C regi
74	353	13.1	343	2 S25644	Ig mu chain C regi
75	349	12.9	453	1 MHRH	Ig mu chain C regi
76	345.5	12.8	454	2 C31933	Ig mu chain C regi
77	331.5	12.3	112	2 B30503	Ig gamma-2a chain
78	325.5	12.0	577	2 I50731	Ig heavy chain - n
79	321.5	11.9	433	2 S31436	Ig epsilon chain -
80	316.5	11.7	240	2 A39016	T-cell surface gly
81	312	11.5	438	1 HVRK2	Ig mu chain C regi
82	308.5	11.4	474	2 I50830	Ig mu chain - lepl
83	305.5	11.3	99	2 S21461	T-cell surface gly
84	301.5	11.2	461	1 HVRKCO	Ig mu chain C regi
85	301.5	11.2	585	2 A46507	Ig alpha chain - c
86	301.5	11.2	684	2 S60266	novel antigen rece
87	300	11.1	342	2 I47175	Ig alpha chain C r
88	297	11.0	339	2 S09264	Ig alpha chain C r
89	296	11.0	357	2 S09269	Ig alpha chain C r
90	296	11.0	367	1 MHRH	Ig mu chain C regi
91	296	11.0	454	2 A46532	Ig mu chain C regi
92	295.5	10.9	343	2 S09272	Ig alpha chain C r
93	294.5	10.9	340	2 B22360	Ig alpha-2 chain C
94	294	10.9	438	1 HVRKCS	Ig mu chain C regi
95	293	10.8	299	1 AHRB	Ig alpha chain C r
96	292	10.8	352	2 S09276	Ig alpha chain C r
97	292	10.8	352	2 S09266	Ig alpha chain C r
98	290	10.7	568	2 A34891	Ig heavy chain pre
99	289	10.7	347	2 S09265	Ig alpha chain C r
100	286.5	10.6	350	2 I56230	Ig alpha-2 chain -
101	286.5	10.6	370	1 HVRKCS	Ig mu chain C regi
102	285.5	10.6	348	2 S09270	Ig alpha chain C r

103	284	10.5	244	2	S12338	Ig heavy chain C r
104	282.5	10.5	348	2	S09273	Ig alpha chain C r
105	282	10.4	347	2	S09274	Ig alpha chain C r
106	281	10.4	393	3	HVRKCI	Ig mu chain C reg1
107	280.5	10.4	99	2	S21462	T-cell surface gly
108	279	10.3	352	2	S05500	Ig alpha-1 chain C
109	277.5	10.3	353	1	A1HU	Ig alpha-1 chain C
110	277	10.3	340	1	A2HU	Ig alpha-1 chain C
111	276.5	10.2	568	2	A45804	Ig gamma-2 chain C
112	275.5	10.2	88	2	A30503	Ig mu chain C reg1
113	275	10.2	357	2	S09267	Ig gamma-2b chain
114	268.5	9.9	350	2	S09271	Ig alpha chain C r
115	264.5	9.8	344	1	AHMS	Ig alpha chain C r
116	263.5	9.8	358	2	S09268	Ig alpha chain C r
117	258	9.5	580	2	A46538	Ig heavy chain C r
118	257.5	9.5	342	2	A43966	Ig heavy chain, see
119	257	9.5	342	2	S43348	Ig epsilon chain C r
120	253.5	9.4	573	2	S12858	Ig mu chain precur
121	245	9.1	335	2	S09275	Ig alpha chain C r
122	243.5	9.0	220	2	C22360	Ig alpha-2 chain C
123	241.5	8.9	448	2	A46503	Ig heavy chain C r
124	223.5	8.3	402	2	S20002	Ig heavy chain, see
125	223.5	8.3	1005	2	T18537	Ig heavy chain - C

## ALIGNMENTS

## RESULT 1

Ig gamma-1 chain C region - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Jan-1991 #sequence revision 18-Aug-1992 #text change 16-Jul-1999  
A:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146  
R:Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Nucleic Acids Res. 10, 4071-4079, 1982  
A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.  
A:Accession: A93433; MUID:82274238; PMID:6287432  
A:Molecule type: DNA  
A:Residues: 1-330 <ELL>  
A:Cross-references: EMBL:Z17370  
A:Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) markers,  
R:Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Submitted to the EMBL Data Library, October 1992  
A:Reference number: S33904  
A:Accession: S36861  
A:Molecule type: DNA  
A:Residues: 2-330 <HAR>  
A:Cross-references: EMBL:Z17370  
R:Takehashi, N.; Ueda, S.; Obata, M.; Nikaide, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of a  
A:Reference number: S33887; MUID:8501943; PMID:6811139  
A:Accession: S33887  
A:Molecule type: DNA  
A:Residues: 88-113/235-330 <TAK>  
A:Cross-references: EMBL:Z17370  
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,  
Biochemistry 9, 3161-3170, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequence  
A:Reference number: A90563; MUID:71064024; PMID:5489771  
A:Accession: B90563  
A:Contents: myeloma protein Eu  
A:Molecule type: protein  
A:Residues: 1-96, R', 98-135 <GUN>  
A:Note: this sequence has the Gln(3) marker, 97-Arg  
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequence  
A:Reference number: A90564; MUID:71064025; PMID:5530842  
A:Contents: Eu

[illegible]

Db 240 ELTKNQVSLTCLVKGFYPSDIAVEMESNQGENNYKTPPVLDSDGSFFLYSKLTIVDKSR 299

Qy 402 WQGNVFGSCVMHEALHNNHYTKSLSLSPG 431

Db 300 WQGNVFGSCVMHEALHNNHYTKSLSLSPG 329

## RESULT 2

569339

Ig heavy chain V region precursor - human

CISpecies: Homo sapiens (man)  
CISDate: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000  
CISAccession: S69339; S72664  
R:Khamlich, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.  
Eur. J. Biochem. 229, 54-60, 1995

A>Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A:Reference number: S69339; PMID:95262687; PMID:7744049

A:Accession: S69339  
A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <KHA>

A:Cross-references: EMBL:X81695

R:Khamlich, A.A.

submitted to the EMBL Data Library, September 1994

A:Reference number: S72664

A:Accession: S72664

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140, 'C', 142-374 <KH2>

A:Cross-references: EMBL:X81695

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 45.2%; Score 1221; DB 2; Length 374;  
Best Local Similarity 71.4%; Pred. No. 2e-65;  
Matches 242; Conservative 19; Mismatches 34; Indels 44; Gaps 5;

Qy 137 QSLTTLT-----LESPGSS-----SPSVQCRSPGKNI 163

Db 35 QTLTLCITSSGSLSGSGVGVGKIRPPQALMLALITWDDDKRSPSLRTRLTITTKOT 94

Qy 164 QGKGLTSLVSOELQDSGTMTCTVL-----QNKVVEFKIDIVPCAPAPKSCDKTHTC- 216

Db 95 SKNQVLTMTNVDPAADATATYYGSGYGGGRFHSWQGLTLYTVSSSPKCDKTHTC 154

Qy 217 ----PELLGSPSVFLFPPPKDTLMSRTPEVTCVVVDVSHEDPEVKFMYVDGVEVNA 272

Db 155 PCPAPPELLGSPSVFLFPPPKDTLMSRTPEVTCVVVDVSHEDPEVKFMYVDGVEVNA 214

Qy 273 KTKPREQVNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEKTSKAGQPREPQ 332

Db 215 KTKPREQVNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEKTSKAGQPREPQ 274

Qy 333 VTTLPSPRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGENNYKTPPVLDSDGSFFLY 392

Db 275 VTTLPSPRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGENNYKTPPVLDSDGSFFLY 334

Qy 393 SKLTVDKSRWQGNVFGSCVMHEALHNNHYTKSLSLSPG 431

Db 335 SKLTVDKSRWQGNVFGSCVMHEALHNNHYTKSLSLSPG 373

## RESULT 3

531866

Ig gamma-1 chain C region - synthetic

CISpecies: synthetic

A>Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

CISDate: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000

CISAccession: S31866

R:Filipula, D.

submitted to the EMBL Data Library, February 1993

A>Description: Screening method for protein-protein interactions of cloned gene products.

A:Reference number: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FLI>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 44.5%; Score 1203.5; DB 4; Length 255;  
Best Local Similarity 97.4%; Pred. No. 1.4e-64;  
Matches 225; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy 206 EPKSCDKTHTC-----PELLGSPSVFLFPPPKDTLMSRTPEVTCVVVDVSHEDPEVK 260

Db 24 ESKSCDKTHTCPCAPPELLGSPSVFLFPPPKDTLMSRTPEVTCVVVDVSHEDPEVK 63

Qy 261 NMYVDGVEVNAKTPREBOYNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEKT 320

Db 84 NMYVDGVEVNAKTPREBOYNSTYRVVSVLTVLHODWLNGEKYCKVSNKALPAPLEKT 143

Qy 321 ISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGENNYKTP 380

Db 144 ISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNQGENNYKTP 203

Qy 381 PVLDSGSFFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTKSLSLSPG 431

Db 204 PVLDSGSFFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTKSLSLSPG 254

## RESULT 4

PT0207

Ig gamma chain C region - chimpanzee

CISpecies: Pan troglodytes (chimpanzee)

CISDate: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999

CISAccession: PT0207

R:Ethlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol. Immunol. 28, 319-322, 1991

A>Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; PMID:91287716; PMID:2062315

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 43.4%; Score 1173.5; DB 2; Length 234;  
Best Local Similarity 92.5%; Pred. No. 7.6e-63;  
Matches 221; Conservative 3; Mismatches 2; Indels 13; Gaps 2;

Qy 192 KVEFKIDIVPCAPAPKSCDKTHTC-----PELLGSPSVFLFPPPKDTLMSRTPEVTC 246

Db 4 KVDKIV-----EPKSCDITHTCPCAAPPELLGSPSVFLFPPPKDTLMSRTPEVTC 55

Qy 247 VVVDVSHEDPEVKFMYVDGVEVNAKTPREBOYNSTYRVVSVLTVLHODWLNGEKYCK 306

Db 56 VVVDVSHEDPEVKFMYVDGVEVNAKTPREBOYNSTYRVVSVLTVLHODWLNGEKYCK 115

Qy 307 KVSNNKALPAPLEKTSKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEM 366

Db 116 KVSNNKALPAPLEKTSKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEM 175

Qy 367 ESNQGENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTKS 425

Db 176 ESNQGENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFGSCVMHEALHNNHYTKS 234

## RESULT 5

A25511

Ig gamma-3 chain C region (allotype G3m(b)) - human

CISpecies: Homo sapiens (man)

CISDate: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999

CISAccession: A25511

R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.



A>Note: this sequence may represent an allelic form or another gamma chain subclass  
C:Comment: The heavy chain disease protein Wis is shown.

C:Genetics:

A:Gene: GDB:IGHG3

A:Cross-references: GDB:119339; OMIM:147120

A:Map position: 14q32.33-14q32.33

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid

F:203-270/Domain: immunoglobulin homology <IMM>

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 41.9%; Score 1133; DB 1; Length 289;  
Best Local Similarity 88.2%; Pred. No. 2,5e-60;  
Matches 209; Conservative 13; Mismatches 9; Indels 6; Gaps 2;

```
QY 201 PCP-APERPSCDKTHTC-----PELIGPSVFLPPPKDITMISTPEVTCVVVDSHE 254
DB 53 PCPRCPERPSCDTPPCPCPCPAPELLIGPSVFLPPPKDITMISTPEVTCVVVDSHE 112
QY 255 DEVRKNNVVDGVEVNNATKPREQVNSTYRVSVTLTLDQMLNGKRYCKVSNKALP 314
DB 113 DEVRKNNVVDGVEVNNATKPREQVNSTYRVSVTLTLDQMLNGKRYCKVSNKALP 172
QY 315 APIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTCLVKGFPSDIAVEMESNQPN 374
DB 173 APIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTCLVKGFPSDIAVEMESNQPN 232
QY 375 NKTTPPVLDSDGSFPLYSKLTVDKSRMQQGVNFCSSVMEALHNHYTKSLSLSPG 431
DB 233 NKTTPPVLDSDGSFPLYSKLTVDKSRMQQGVNFCSSVMEALHNHYTKSLSLSPG 289
```

#### RESULT 8

IG gamma-2 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 30-Apr-1981 #sequence revision 13-Jun-1983 #text\_change 21-Jul-2000

C:Accession: A93906; A92809; A90752; A93132; A02148

R:Ellison, J.; Hood, L.

Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982

A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con

A:Reference number: A93906; MUID:82197621; PMID:6804948

A:Accession: A93906

A:Molecule type: DNA

A:Residues: 1-326 <ELI>

A:Cross-references: GB:V00554; GB:J00230; NID:932759; PIDN:CABS8438.1; PID:9606056

A>Note: Lys-326 is probably removed posttranslationally

R:Wang, A.C.; Tung, E.; Fudenberg, H.H.

J. Immunol. 125, 1048-1054, 1980

A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and E

A:Reference number: A92809; MUID:81007873; PMID:6774012

A:Accession: A92809

A:Molecule type: protein

A:Residues: 1-119, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <MAN>

A>Note: Trp-156 is at or near the complement-binding site

R:Connell, G.E.; Parr, D.M.; Hofmann, T.

Can. J. Biochem. 57, 758-767, 1979

A:Title: The amino acid sequences of the three heavy chain constant region domains of a

A:Reference number: A90752; MUID:80001357; PMID:113060

A:Accession: A90752

A:Molecule type: protein

A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-

A>Note: this sequence has since been revised

R:Hofmann, T.; Parr, D.M.

Mol. Immunol. 16, 923-925, 1979

A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g

A:Reference number: A93132; MUID:80114419; PMID:118920

A:Accession: A93132

A:Molecule type: protein

A:Residues: 238-275 <HOF>  
R:Hofmann, T.; Parr, D.M.  
submitted to the Atlas, March 1980  
A:Reference number: A94591  
A:Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268  
A>Note: the revised sequence differs from that shown in having 60-Ala and in the amidat

R:Milstein, C.; Frangione, B.

Biochem. J. 121, 217-225, 1971

A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.

A:Reference number: A90253; MUID:72033500; PMID:4940472

A:Contents: annotation; myeloma protein Sa, disulfide bonds

R:Frangione, B.; Milstein, C.; Pink, J.R.L.

Nature 221, 145-148, 1969

A:Title: Structural studies of immunoglobulin G.

A:Reference number: A93157; MUID:69064124; PMID:5782707

A:Contents: annotation; Sa, disulfide bonds

C:Genetics:

A:Gene: GDB:IGHG2

A:Cross-references: GDB:119338; OMIM:147110

A:Map position: 14q32.33-14q32.33

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kai

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1:

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:133-202/Domain: immunoglobulin homology <IM1>

F:239-306/Domain: immunoglobulin homology <IM2>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:27-93,140-200,246-304/Disulfide bonds: #status experimental

F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental

F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.8%; Score 1130.5; DB 1; Length 326;  
Best Local Similarity 82.6%; Pred. No. 4e-60;  
Matches 214; Conservative 12; Mismatches 12; Indels 21; Gaps 3;

```
QY 181 TWTCV-----LQNKVFEK-IDIVCPAPERPSCDKTHTCBLGSPVFLPPPKP 232
DB 80 ITTCVHKPKSNTKDKVTKKVCCEPCPCP-----VAGPSVFLPPPKP 126
QY 233 KDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKTKPREEQNSTYRVSVLT 292
DB 127 KDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKTKPREEQNSTYRVSVLT 186
QY 293 VTHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTC 352
DB 187 VTHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLPPSRDELTKNQVSLTC 246
QY 353 LVKGFYPSDIAVEMESNQPNPKYTKTTPVLDSDGSFPLYSKLTVDKSRMQQGVNFCSCV 412
DB 247 LVKGFYPSDIAVEMESNQPNPKYTKTTPVLDSDGSFPLYSKLTVDKSRMQQGVNFCSCV 306
QY 413 MHEALHNHYTKSLSLSPG 431
DB 307 MHEALHNHYTKSLSLSPG 325
```

#### RESULT 9

GAHU

IG gamma-4 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 02-Apr-1982 #sequence revision 02-Apr-1982 #text\_change 16-Jul-1999

C:Accession: A90933; A90249; A02150

R:Ellison, J.; Buxbaum, J.; Hood, L.

DNA 1, 11-18, 1981

A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.

A:Reference number: A90933; MUID:83157104; PMID:6299662

A:Accession: A90933

A:Molecule type: DNA

A:Residues: 1-327 <ELI>

A>Note: the sequence was determined from the germline gene

R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.

A:Accession: A32722  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 26-426,428-458 <C>  
R:Garr, S.A.; Hemling, M.E.; Polena-Wasserman, G.; Sweet, R.W.; Anumula, K.; Barr, J.R.; J. Biol. Chem. 264, 21286-21295, 1989  
A:Title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 receptor  
A:Reference number: A34194, MUID:90078223, PMID:252374  
A:Contents: disulfide bonds; carbohydrate-binding sites  
A:Accession: A34194  
A:Molecule type: Protein  
A:Residues: 26-394 <C>  
R:Lederman, S.; Demattino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, P.  
Mol. Immunol. 28, 1171-1181, 1991  
A:Title: A single amino acid substitution in a common African allele of the CD4 molecule  
A:Reference number: A53287; MUID:92072595; PMID:1961196  
A:Accession: A53287  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 250-264,'W',266-280 <L>  
A:Note: sequence extracted from NCBI backbone (NCBIP:68249)  
R:Edwards, M.C.; Gibbs, R.A.  
Genomics 14, 590-597, 1992  
A:Title: A human dimorphism resulting from loss of an Alu  
A:Reference number: I54176; MUID:93052387; PMID:1330888  
A:Accession: I54176  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-72 <R>  
A:Cross-references: GB:I47924; GB:M86525; GB:U22506; NID:G1633547; PIDN:AA51309.1; PID  
R:Hodge, T.W.; Saeko, D.R.; McConaughy, J.S.  
Hum. Immunol. 30, 99-104, 1991  
A:Title: Humans with OKT4-epitope deficiency have a single nucleotide base change in the  
A:Reference number: I54297; MUID:91216786; PMID:1708753  
A:Accession: I54297  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-264,'W',266-458 <R>  
A:Cross-references: GB:M35160; NID:G179143; PIDN:AAA16069.1; PID:G179144  
C:Comment: Macrophage tropic strains of HIV-1 bind to a complex of chemokine (C-C) receptor  
C:Genetics:  
A:Gene: GDB:CD4  
A:Cross-references: GDB:I19767; OMIM:186940  
A:Map position: 12pter-12p12  
A:Introns: 16/3  
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
C:Keywords: AIDS; duplication; glycoprotein; T-cell; transmembrane protein  
F:1-35/Domain: signal sequence #status predicted <SIG>  
F:26-458/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>  
F:34-111/Domain: immunoglobulin homology <IM1>  
F:136-186/Domain: immunoglobulin homology #status atypical <IM2>  
F:216-299/Domain: immunoglobulin homology <IM3>  
F:321-372/Domain: immunoglobulin homology <IM4>  
F:397-420/Domain: transmembrane #status predicted <TM>  
F:421-458/Domain: intracellular #status predicted <INT>  
F:41-109,155-184,328-370/Disulfide bonds: #status experimental  
F:296,325/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match	37.9%	Score 1023	DB 1	Length 458
Best Local Similarity	99.5%	Pred. No. 1.5e-53		
Matches 199	Conservative 0	Mismatches 1	Indels 0	Gaps 0

QY	1	MNRGVPFRHLLLVQALALPAATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNSNQIK	60
DB <td>1</td> <td>MNRGVPFRHLLLVQALALPAATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNSNQIK</td> <td>60</td>	1	MNRGVPFRHLLLVQALALPAATQGNKVVLGKGGDTVELTCTASQKSIQPHMKNNSNQIK	60

QY	61	ILNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLTIEDSDTYICEVEDQKEEVQL	120
DB <td>61<th>ILNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLTIEDSDTYICEVEDQKEEVQL</th><th>120</th></td>	61 <th>ILNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLTIEDSDTYICEVEDQKEEVQL</th> <th>120</th>	ILNQGSFLTKGSPKLNDRADSRSLWDQGNFPLIINKLTIEDSDTYICEVEDQKEEVQL	120

QY	121	LVFELTANSTHLLQGSFLTLTLESPGSSPSVQCRPRGNIGQKTLISVQLELDQSG	180
DB <th>121</th> <th>LVFELTANSTHLLQGSFLTLTLESPGSSPSVQCRPRGNIGQKTLISVQLELDQSG</th> <th>180</th>	121	LVFELTANSTHLLQGSFLTLTLESPGSSPSVQCRPRGNIGQKTLISVQLELDQSG	180

QY 181 TWCTVLONQKVEFKIDIV 200  
 |||||  
 DB 181 TWCTVLONQKVEFKIDIV 200

## RESULT 11

G3MSM  
 Ig gamma-3 chain C region, membrane-bound form - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 13-Aug-1986 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999  
 C/Accession: A02156; A02155  
 R/Author: J.A.; Word, C.J.; Rimm, D.; Der-Balan, G.P.; Martinez, H.M.; Tucker, P.W.; Blat  
 EMO J. 3, 2041-2046, 1984  
 A/Title: Structure analysis of the murine IgG3 constant region gene.  
 A/Reference number: A02156; MUID:85027161; PMID:6092053  
 A/Accession: A02156  
 A/Molecule type: DNA  
 A/Residues: 1-398 <MEL>  
 A/Cross-references: GB:J00451; NID:g194392; PIDN:AAB59655.1; PID:g194433  
 A/Note: the sequence was determined from the germline gene  
 R/Komarov, M.; Clayton, L.; Rogers, J.; Robertson, S.; Kettman, J.; Wall, R.  
 Nucleic Acids Res. 11, 6775-6785, 1983  
 A/Title: The structure of the mouse immunoglobulin in gamma-3 membrane gene segment.  
 A/Reference number: A02155; MUID:84041483; PMID:6314258  
 A/Accession: A02155  
 A/Molecule type: DNA  
 A/Residues: 328-332, 'G', 334-341, 'Q', 343-387, 'F', 389-398 <KOM>  
 A/Cross-references: GB:K00688  
 A/Note: the sequence was determined from the germline gene  
 C/Genetics:  
 A/Intons: 97/1; 113/1; 223/1; 328/1; 371/3  
 C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (L)  
 chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C/superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: alternative splicing; duplication; glycoprotein; heterotrimer; immunoglob  
 F/19-83/Domain: immunoglobulin homology <IM1>  
 F/97-112/Region: hinge  
 F/136-205/Domain: immunoglobulin homology <IM2>  
 F/242-309/Domain: immunoglobulin homology <IM3>  
 F/346-362/Domain: transmembrane #status predicted <TM>  
 F/363-398/Domain: intracellular #status predicted <INT>  
 F/179,322/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Query Match

35.4%; Score 955.5; DB 1; Length 398;  
 Best Local Similarity 61.4%; Pred. No. 1,3e-49;  
 Matches 180; Conservative 40; Mismatches 58; Indels 15; Gaps 4;

QY 170 SVSLELDOSGTV-----TCTVLONQKVEFKIDIV---VPCAPPEPKSCDKHTCP--EL 219  
 |||  
 DB 62 SLSSLVTVSSTSPSCQVTCNVNAPASKTELKRIEPRIPKSPSTPGS-----SCPPGNI 116  
 QY 220 LGSPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNNYVGVGVNNAKTRPREE 279  
 |||||  
 DB 117 LGSPSVFIFPPKPKDMLMISLTPEKTCVVVDVSEDDPDVAVSVFVNKKEHTAMTPREA 176  
 QY 280 QVNSTTRVSVTLVLDQMLNGKEYCKRKNKLLPARIETKISKAGQPREPVYTLPPS 339  
 |||||  
 DB 177 QVNSTTRVSVTLVLDQMLNGKEYCKRKNKLLPARIETKISKAGQPREPVYTLPPS 236  
 QY 340 RDELTKNOVSLTLGVGVFVPSDIAVEMESNGDEPNYKTPPYLSDGSEFLLSKLTIVK 399  
 |||||  
 DB 237 RDEMSKKRSLTLGVGVFVPSDIAVEMESNGDEPNYKTPPYLSDGSEFLLSKLTIVK 296  
 QY 400 SRNQGNVFSQVMEALNHYTQKSLSPGLDDETCADQDGLDGLMTT 452  
 |||||  
 DB 297 DSVLQGEIFTCQVVEALNHYTQKSLSPPELEMETCAEQDGLDGLMTT 349

## RESULT 12

G1MSM  
 Ig gamma-1 chain C region, membrane-bound form - mouse  
 C/Species: Mus musculus (house mouse)

C/Date: 17-Dec-1982 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999  
 C/Accession: B02159; A02160; B02158  
 R/Honjo, T.; Obata, M.; Yamawaki-Kataoka, Y.; Kataoka, T.; Kawakami, T.; Takahashi, N.;  
 Cell 18, 559-568, 1979  
 A/Title: Cloning and complete nucleotide sequence of mouse immunoglobulin gamma1 chain  
 A/Reference number: A02159; MUID:80045036; PMID:115593  
 A/Accession: B02159  
 A/Molecule type: DNA  
 A/Residues: 1-393 <HON>  
 A/Cross-references: GB:J00453  
 A/Note: the sequence was determined from the germline gene  
 R/Tyler, B.M.; Cowan, A.F.; Gerondakis, S.D.; Adams, J.M.; Bernard, O.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 2008-2012, 1982  
 A/Title: mRNA for surface immunoglobulin gamma chains encodes a highly conserved transm  
 A/Reference number: A02160; MUID:82197626; PMID:6804950  
 A/Accession: A02160  
 A/Molecule type: mRNA  
 A/Residues: 323-393 <TYL>  
 R/Rogers, J.; Choi, E.; Souza, L.; Carter, C.; Word, C.; Kuehl, M.; Eisenberg, D.; Wall,  
 Cell 26, 19-27, 1981  
 A/Title: Gene segments encoding transmembrane carboxyl terminus of immunoglobulin gamma  
 A/Reference number: A02158; MUID:82115295; PMID:6799207  
 A/Accession: B02158  
 A/Molecule type: DNA  
 A/Residues: 323-366 <ROG>  
 A/Note: this sequence is the translation of the first exon of the M segment  
 C/Comment: Cell lines producing IgG contain two mRNA species for Ig gamma chains. The m  
 ncode membrane-bound chains in that it contains an alternative 3' end, encoded in separa  
 C/Genetics:  
 A/Intons: 1/1; 98/1; 111/1; 218/1; 323/1; 366/3  
 C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (L)  
 chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
 C/superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: alternative splicing; duplication; glycoprotein; heterotrimer; immunoglob  
 F/131-200/Domain: immunoglobulin homology <IM1>  
 F/340-357/Domain: transmembrane #status predicted <TM>  
 F/358-393/Domain: intracellular #status predicted <INT>  
 F/174,278/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Query Match

34.5%; Score 932.5; DB 1; Length 393;  
 Best Local Similarity 55.0%; Pred. No. 2,9e-48;  
 Matches 177; Conservative 50; Mismatches 60; Indels 35; Gaps 5;

QY 133 LQGSQSLTLTSPSSPSVOCSPRGKNIQGGKTLVSQLELDOSGTVTLQ--NQ 190  
 |||||  
 DB 56 VQGSQSLTLTSSSVTVSSP-----RPSFTVCNVNAPASS 90  
 QY 191 KKYVEFKIDIVPCAPPEPKSCDKHTCPPELLGSPVFLPPPKDPTLMISRTPEVTCVVVD 250  
 |||||  
 DB 91 TVYDKKIVRDGCG-KPCIC-----TVPEV---SVFIFPPKPKDPTLTITLTKVTCVVVD 142  
 QY 251 VSHEDPEVKFNNYVGVGVNNAKTRPREEQVNSTYRVVSVTLVLDQMLNGKEYCKKCVSN 310  
 |||||  
 DB 143 ISKDPEVQFSWFDVDEVTQOTQPREQFNSSTRESVELPIHMDQMLNGKEYCKRVNS 202  
 QY 311 KALPARIETKISKAGQPREPVYTLPPSRDRLTKNOVSLTLGVGFYSDIAVEMESNG 370  
 |||||  
 DB 203 AAFPAPIETKISKAGQPREPVYTLPPSRDRLTKNOVSLTLGVGFYSDIAVEMESNG 262  
 QY 371 QENNYKTPPYLSDGSEFLLSKLTIVDSRNQGNVFSQVMEALNHYTQKSLSP 430  
 |||||  
 DB 263 QPAENYKTPPYLSDGSEFLLSKLTIVDSRNQGNVFSQVMEALNHYTQKSLSP 322  
 QY 431 GLDDETCADQDGLDGLMTT 452  
 |||||  
 DB 323 GLDDETCADQDGLDGLMTT 344

## RESULT 13

S32080  
 Ig heavy chain precursor (B/MT.4A.17.H5.A5) - bovine  
 N/Alternate names: Ig gamma-1 chain C region (clone 8.10)  
 C/Species: Bos primigenius taurus (cattle)



C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
 C/Accession: S22080; S06610; A31303  
 R/Sanders, P.G.  
 submitted to the EMBL Data Library, November 1991  
 A/Reference number: S22080  
 A/Accession: S22080  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-470 <SAB>  
 A/Cross-references: EMBL:X62916; NID:q439; PIDN:CAA44699.1; PID:q440  
 R/Symons, D.B.A.; Clarkson, C.A.; Beale, D.  
 Mol. Immunol. 26, 841-850, 1989  
 A/Title: Structure of bovine immunoglobulin constant region heavy chain gamma 1 and gamma 2  
 A/Reference number: S06610; MUID:90097956; PMID:2513487  
 A/Accession: S06610  
 A/Molecule type: DNA  
 A/Residues: 142-470 <SYM>  
 A/Cross-references: EMBL:X16701  
 A/Note: The sequence was determined from the germline gene  
 C/Genetics:  
 A/Genes: Ig CH gamma-1  
 A/Introns: 98/1; 111/1; 221/1  
 C/Superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: glycoprotein; heterotetramer; immunoglobulin; membrane protein  
 F/161-225/Domain: immunoglobulin homology <IMM>  
 F/118/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.9%; Score 915.5; DB 2; Length 470;  
 Best Local Similarity 43.6%; Pred. No. 3.7e-47;  
 Matches 214; Conservative 52; Mismatches 136; Indels 89; Gaps 15;

QY 6 PFRHLTLVQ-----LALLPATQGNKVLKSGDVTCTCTAS--CKSKIQFMKNSNQ 58  
 DB 3 PLMTLLFVLSAPIGVLSQVQRESGPLV--KPSQTLSLTCTVSGFSLVALTMVRAP 60  
 QY 59 IKILGNQGSFLTKGSKLANDRDSRRSLMDQCN--FLIIRKMLKIEPSDYIC----- 109  
 DB 61 GKALEWVGITSGGTYYNPAKSLSTTKENSKSQVLSVSVTPBEDTATYCCARSTYG 120  
 QY 110 EVEDQKEVQLVFGLTANSPTHLIQGSLTLTLESPPSGSPSYQ-----CRSPRGKNIQ 164  
 DB 121 EVGD-----GALIDA-----WGQGLVTVSSASTAPKYPYPLSSCCGDKSSSTVT 165  
 QY 165 GSKTSLVSQLE-----LQDSG-----TWLCTV 186  
 DB 166 LGCLVSSVMPPEPVYTNNSGALKSGVHTFPAVLQSSGLYSLSMTVTCGTSQGTFCNV 225  
 QY 187 LQ--NOKKVEFKIDIVPCPAPBPYKSCDKTHTC--PELIGPSVFLFPKPKDTLMISRT 242  
 DB 226 AHPASSTVQDAVD---PTCKPSPCD---CCPPELPFGSPSVFFPKPKPDTLTISGTP 278  
 QY 243 EVTCVVVDVSHEDPEVKFNWYVDGVEFNNAKTKRPEQYNSTYVSVLTALHDDMLNGK 302  
 DB 279 EVTCVVVDVGHDDPEVKFNSWFDVDEVTATTKPEEDFNSITRVASLRIQHDQDWGKG 338  
 QY 303 EYCKVNSKALPAPIEKTISRAKQPRPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 362  
 DB 339 EFKCKVNHGELPAPIVRTISRTKGPAPRPQVYVLAPEELSKSTVSLTCLVNTSFYPIY 398  
 QY 363 AVENESNQP--ENNYKTPPVLDSDGSFYLKSLTVDKSRMQQGNVFSGVMEBALHNN 420  
 DB 399 AVENQORNQPESEDEKYGTPPOLDADSSYFLYSKLRVDRNSMQEGDVTYTCVVMEDALHNN 458  
 QY 421 YTKSLSLSPG 431  
 DB 459 YTKSTSKSAG 469

RESULT 14  
 GZMSAM  
 Ig gamma-2a chain C region, membrane-bound form - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 19-Feb-1984 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999

C/Accession: A02154; B32657; I57809  
 R/Yanawaki, Katsoka, Y.; Nakai, S.; Miyata, T.; Honjo, T.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 2623-2627, 1982  
 A/Title: Nucleotide sequences of gene segments encoding membrane domains of immunoglobulin  
 A/Reference number: A02154; MUID:82221190; PMID:6283537  
 A/Accession: A02154  
 A/Molecule type: DNA  
 A/Residues: 329-399 <YAM>  
 A/Cross-references: GB:J00471  
 A/Note: The sequence was determined from the germline gene  
 R/Yanawaki, Katsoka, Y.; Miyata, T.; Honjo, T.  
 Nucleic Acids Res. 9, 1365-1381, 1981  
 A/Title: The complete nucleotide sequence of mouse immunoglobulin gamma 2 A gene and epsilon  
 A/Reference number: A32657; MUID:81198976; PMID:6262729  
 A/Accession: B32657  
 A/Molecule type: DNA  
 A/Residues: 1-329 'K' <YAN>  
 R/Hall, B.; Micalarek, C.  
 Mol. Immunol. 26, 819-826, 1989  
 A/Title: Sequence and polyadenylation site determination of the murine immunoglobulin gamma  
 A/Reference number: I57809; MUID:90097953; PMID:2513466  
 A/Accession: I57809  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 333-399 <RES>  
 A/Cross-references: GB:M35032; NID:q194478; PIDN:AAA37919.1; PID:q387217  
 C/Comment: The sequence of residues 1-328 was assumed to be identical with the corresponding  
 C/Genetics: Cell lines producing IgG contain two mRNA species for Ig gamma chains. The major  
 C/Genetics: It contains an alternative 3' end, encoded in separate exons, that is homologous with  
 A/Introns: 1/1; 98/1; 114/1; 224/1; 329/1; 372/1  
 C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (Kap  
 C/Superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin  
 F/137-206/Domain: immunoglobulin homology <IMM>  
 F/346-393/Domain: transmembrane #status predicted <TM>  
 F/364-399/Domain: intracellular #status predicted <INT>  
 F/180/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.5%; Score 906.5; DB 1; Length 399;  
 Best Local Similarity 60.5%; Pred. No. 1e-46;  
 Matches 173; Conservative 34; Mismatches 60; Indels 19; Gaps 5;

QY 179 SGTW-----TCTVQ--NOKKVEFKID-----IYCPAPBPYKSCDKTHTCELGSPSVF 226  
 DB 72 SSTWPSQSTICNVNAPHSSTKVDKIEPRQPTIFCP---PKC-----PAPNLLGSPSVF 124  
 QY 227 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEFNNAKTKRPEQYNSTYR 286  
 DB 125 IFPKKIDVLMISLSPVTCVVVDVSEDDPDVQSWFNVNVEVHTAQOTHRBEDYNSTLR 184  
 QY 287 VVSIVLTALHDDMLNGKVKCKVNSKALPAPIEKTISRAKQPRPQVYTLPPSRDELTKN 346  
 DB 185 VVSALPIQHDQWMSGKFKKVNKKDLPAPIERTISPKKSVAPQVYVLPPEEEMTKK 244  
 QY 347 QVSLTCLVKGFYPSDIAVENESNQPENNYKTPPVLDSDGSFYLKSLTVDKSRMQQGN 406  
 DB 245 QVTLTCNVITPMPEDIVETNTNGKTELANTKMTEPVLDSGSYFMYSKLRVEKNWERN 304  
 QY 407 VFGSGVMEALHNNYTKSLSPGLQLDCTCAEADQDELGLMTT 452  
 DB 305 SYSQSVVHEGLHNNHTTKSFSTRPGDLDDVCAEAQDGLGLMTT 350

RESULT 15  
 I47159  
 Ig gamma 2a chain constant region - pig (fragment)  
 C/Species: Sus scrofa domestica (domestic pig)  
 C/Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
 C/Accession: I47159  
 R/Kacskovics, I.; Sun, J.; Butler, J.E.  
 J. Immunol. 153, 3565-3573, 1994

A>Title: Five active subclasses of swine Ig identified from the cDNA sequences of a  
A:Reference number: 147158, MUID:95015845; PMID:7930579  
A:Accession: 147159  
A:Strctus: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Crs98-references: EMBL:U03779; NID:g433123; PIDD:AA52217.1; PID:g433124  
C:Genetics:  
A:Gene: IGC2a  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match	33.3%;	Score 901;	DB 2;	Length 328;
Best Local Similarity	58.5%;	Pred. No. 1.7e-46;		
Matches 182;	Conservative 38;	Mismatches 57;	Indels 34;	Gaps 9

```

0Y 137 0SLTLTLESPRSS-----PSVQCRSPRG-KNIQGGKTLTSVQLELODSESTCTYLQNG 190
Db 35 EPLVATWNSGALSLSGVHTPSPV--LQPSGLTSLSSWVTLPASYL---SSKSYLCNN--NH 87
0Y 191 KAVEFKID-----IVPCPAPEPKSCDTHNCTPELLG--GVSFLFPFKKOTLMTSRTP 242
Db 88 PATTTKVDKRGVGTKTTRPCP-----ICPACESQSPSFIITPPKXOTLMTSRTP 136
0Y 243 EVTCVVVDVSHEDPEVKFMYVDDGVENHNAKTRPREEQYNSITRVVSVLTVLHODMLNGK 302
Db 137 QVTCVVVDVSGNPEVFQFSMYDDGVENHNAQTRPKCEQFNSTYRVVSVLTPLQHODMLNGK 196
0Y 303 EYKCKVSKNALPAPLEKTIISKAKGRPEQVYTLTPRSDELTKNQVSLTCLVKGFYPSDI 362
Db 197 EFCKCKVNNKDLPAPIRLIISKAKGRPEQVYTLTPRHAELSSKVSITCLVIGFYPDI 256
0Y 363 AVENESNGQ--PENNYKTTTPVLDSGSEFFVLSKLTVDVSKRMQGNVPSGVNHEALHNH 420
Db 257 DVEWQRNGRPEPEGNVRRITPPQODVDGTYFLYSKFSVDKASMGWGGIFQCAVNHAEALHNH 316
0Y 421 YTKQSLSLSPG 431
Db 317 YTKQSLSLSPG 327

```

## RESULT 16

SHRB  
Ig gamma chain C region - rabbit  
C.Species: Oryctolagus cuniculus (domestic rabbit)  
C.Date: 24-Apr-1984 #sequence revision 15-Nov-1994 #text change 16-Jul-1999  
C.Accession: A01749; A00290; A03928; A00245; A04416; A02161  
R.Bernslein, K.E.; Alexander, C.B.; Mage, R.G.  
Immunogenetics 18, 387-397, 1983  
A.Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplotype  
A.Reference number: A01749; MUID:84030930; PMID:6313520  
A.Accession: A01749  
A.Molecule type: mRNA  
A.Residues: 1-523 <BER>  
A.Note: This sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Thr  
R.Pratt, D.M.; Mole, L.E.  
Biochem. J. 151, 337-349, 1975  
A.Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin G  
A.Reference number: A00290; MUID:76135469; PMID:1243651  
A.Accession: A00290  
A.Molecule type: Protein  
A.Residues: 1-47, 'E', 49-71, 'FV', 72-128 <PRA>  
R.Martens, C.L.; Moore, K.W.; Steilmetz, M.; Hood, L.; Knight, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
A.Title: Heavy chain genes of rabbit IgG, isolation of a cDNA encoding gamma heavy chain  
A.Reference number: A03928; MUID:83299917; PMID:6195512  
A.Accession: A03928  
A.Molecule type: mRNA  
A.Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
A.Cross-references: GB:M6426; NID:G16511; PIDN:AA03189, 1; PID:G165112  
R.Fritschner, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
Biochem. J. 116, 249-259, 1970

A/Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin A/Reference number: A90245; MUID:70110015; PMID:5461106

A: Molecule type: protein  
A: Residues: 132-143 'E', 145-161 <FRU>  
R: Hill, R.L.; Lebowitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiksell  
A: Reference number: A94416

```
A:Molecule type: protein
A:Residues: 129-131,155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, 'Q',
A>Note: this has the e15 allelopathic marker, 185-191
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-82/Domain: immunoglobulin homology <IM1>
F:130-199/Domain: immunoglobulin homology <IM2>
F:236-303/Domain: immunoglobulin homology <IM3>
F:173/Binding site: carboxylate (Asn) (covalent) #status predicted
```

Query Match	33.3%;	Score 899;	DB 1;	Length 323;
Best Local Similarity	61.0%;	Pred. No. 2.2e-46;		
Matches 178;	Conservative 34;	Mismatches 64;	Indels 16;	Gaps 5

QY 142 TLLESPGSSPVQCSPRGKNIGGKTIISVGLDSDSTWCTCYLQNGKVEFKDIYP 201  
 Db 45 TLTNVRIRFPVSROS-----GLYSLSSVSIVSSSQPTVCNA--HPIATNKD--- 92  
 QY 202 CPAPEPKSCDKHTHC-PELLGSPVFLPPPKKDTLMSRTPEVTCVVVDVSHEDPEYK 259  
 Db 93 -KTVASPTCSKP-TCPEPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSDQDDPEVQ 150  
 QY 260 FNNVYDGVENNAKTKPREBOYNSIVRVVSVLTVLHOMLNKEKYCKSNALPAPIEK 319  
 Db 151 FTWIIINQVTRARPPLEQCPNSITIRVASTILPIIHODLRGKEFECKYHNALPAPIEK 210  
 QY 320 TISAKGQPREPOVYITLPISRDDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKT 379  
 Db 211 TISKARGQPLPEPKYVTMGPPRELLISRSASVSLTCLMGFEPYSDISVEEMNGKAEDNYKTT 270  
 QY 380 PPVLDSDSFLYSLKLTVDKSRMOGQNPVSCVMEHALNHNHTQSLSLSPG 431  
 Db 271 PAVLDSDSFLYKNTLSVPTSEMQGDVFTCSVMHEALNHNHTQSLSLSPG 322

## RESULT 17

14/192  
Ig gamma 4 chain constant region - pig (fragment)  
C|Species: Sus scrofa domestica (domestic pig)  
C|Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C|Accession: I47162  
R|Kackskovic, I.; Sun, J.; Butler, J.E.  
J|Immunol. 153, 3565-3573, 1994  
A|Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
A|Reference number: I47158; MUID:95015845; PMID:7930579  
A|Accession: I47162  
A|Status: preliminary; translated from GB/EMBL/DBD  
A|Molecule type: mRNA  
A|Residues: 1-277 <KAC>  
A|Cross-references: EMBL:U03782; NID:g433129; PIDD:AA52220.1; PID:g433130  
C|Genetic8  
A|Gene: IgG4  
C|Superfamily: immunoglobulin C region; immunoglobulin homology  
F|82-151|Domain: immunoglobulin homology <IM>

```

Query Match      33.2%  Score 998 ; DB 2; Length 277;
      Similarity 62.8%  Pred. No. 2,1e-46;
      Local      172; Conservative 33; Mismatches 43; Indels 26; Gaps 6
Oy 168 TLVSQLELDGSGTCTVLONOKRVEKID-----IVCPADPEPKSCDKHTTCPELL 220
Db 19 IVPASSL---SSKSTACNV--NHRAPTTKRKVDGRGVGKTKRPPC-----ICPAGE 62

```

QY 221 G-GPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 279  
 DB 63 GGPGRAPFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 122  
 QY 280 QYNSTYRVAVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS 339  
 DB 123 QPNSTYRVAVSVLTPLIOHOMLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS 182  
 QY 340 RDELTKNOVSLTCLVKGFPYPSDIAVEMESNGQ--PENNYKTTTPVLDSGSFFLYSKLTV 397  
 DB 183 TEELSRKSVTLTCLVGTGFPYPPDIDVEMQNGQPEPEGNVYRTTPPQDDVGTGYLYSKLAV 242  
 QY 398 DKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPG 431  
 DB 243 DKASWQGRDFTQCAVNHMALHNHYTQKSISLTPG 276

## RESULT 18

147160  
 Ig gamma 2b chain constant region - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
 C:Accession: 147160  
 R:Kacskovics, I.; Sun, J.; Butler, J.E.  
 J:Immunol. 153, 3565-3573, 1994  
 A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
 A:Reference number: 147158; MUID:95015845; PMID:7930579  
 A:Accession: 147160  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Statues: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <KAC>  
 A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AA52218.1; PID:9433126  
 C:Genetics:  
 A:Gene: IgG2b  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:133-202/Domain: immunoglobulin homology <IMW>

Query Match 33.2%; Score 898; DB 2; Length 328;  
 Best Local Similarity 62.8%; Pred. No. 2.6e-46;  
 Matches 172; Conservative 33; Mismatches 43; Indels 26; Gaps 6;

QY 168 TLVSQLELDQSGTWTCTVLQ--NOKKVEFKID-----IVCPAPEPKSCDKHTHCPRL 220  
 DB 70 TVPASSL---SSKSYTCNV--NHPATTKVDKRVGTKTKPCP-----ICPACE 113  
 QY 221 G-GPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 279  
 DB 114 SPGPSVFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 173  
 QY 280 QYNSTYRVAVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS 339  
 DB 174 QPNSTYRVAVSVLTPLIOHOMLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS 233  
 QY 340 RDELTKNOVSLTCLVKGFPYPSDIAVEMESNGQ--PENNYKTTTPVLDSGSFFLYSKLTV 397  
 DB 234 AEELSRKSVTLTCLVGTGFPYPPDIDVEMQNGQPEPEGNVYRTTPPQDDVGTGYLYSKRSV 293  
 QY 398 DKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPG 431  
 DB 294 DKASWQGGIFQCAVNHMALHNHYTQKSISLTPG 327

## RESULT 19

19 gamma-2 chain C region - guinea pig  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 07-May-1981 #sequence\_revision 07-May-1981 #text\_change 16-Jul-1999  
 C:Accession: A94553; A90352; A90359; A90384; A90385; A02151  
 R:Trischmann, T.M.  
 submitted to the Atlas, April 1975  
 A:Reference number: A94553  
 A:Accession: A94553

A:Molecule type: protein  
 A:Residues: 1-3 <TRI>  
 R:Birshtein, B.K.; Hussain, Q.Z.; Cebra, J.J.  
 Biochemistry 10, 18-25, 1971  
 A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am  
 A:Reference number: A90352; MUID:71058471; PMID:5538606  
 A:Accession: A90352  
 A:Molecule type: protein  
 A:Residues: 4-68 <BIR>  
 R:Turner, K.J.; Cebra, J.J.  
 Biochemistry 10, 9-17, 1971  
 A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. Am  
 A:Reference number: A90359; MUID:71058486; PMID:5538616  
 A:Accession: A90359  
 A:Molecule type: protein  
 A:Residues: 69-133,312-329 <TUR>  
 R:Tracey, D.E.; Cebra, J.J.  
 Biochemistry 13, 4796-4803, 1974  
 A:Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.  
 A:Reference number: A90384; MUID:75036073; PMID:4429665  
 A:Accession: A90384  
 A:Molecule type: protein  
 A:Residues: 134-226 <TRA>  
 R:Trischmann, T.M.; Cebra, J.J.  
 Biochemistry 13, 4804-4811, 1974  
 A:Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.  
 A:Reference number: A90385; MUID:75036073; PMID:4609467  
 A:Accession: A90385  
 A:Molecule type: protein  
 A:Residues: 227-311 <TR2>  
 R:Oliveira, B.; Lamu, M.E.  
 Biochemistry 10, 26-31, 1971  
 A:Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.  
 A:Reference number: A90354; MUID:71058474; PMID:4922544  
 A:Contents: annotation; disulfide bonds  
 A>Note: Cys-16 is involved in a heavy-light chain bond  
 A>Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds  
 C:Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.  
 C:Complex: An immunoglobulin heterotrimeric subunit consists of two identical light (kappa  
 chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:21-81/Domain: immunoglobulin homology <IM1>  
 F:135-204/Domain: immunoglobulin homology <IM2>  
 F:241-310/Domain: immunoglobulin homology <IM3>  
 F:28-79/Disulfide bonds: #status experimental  
 F:142-202/Disulfide bonds: #status experimental  
 F:178/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:248-308/Disulfide bonds: #status experimental

Query Match 33.1%; Score 894.5; DB 1; Length 329;  
 Best Local Similarity 62.5%; Pred. No. 4.2e-46;  
 Matches 172; Conservative 34; Mismatches 60; Indels 9; Gaps 4;

QY 163 IOGKTLVSQLELDQSGTWTCTVLQ--NOKKVEFKIDIVCPAPEPKSCDKHTHCPRL 218  
 DB 58 IQGSLVSLTGMVTPSSQKATCVVAHPASSTKVDKTEVPIRTPEPBCCTCK---CPPE 114  
 QY 219 LIGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 278  
 DB 115 NLGGPSVFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE 174  
 QY 279 QYNSTYRVAVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS 338  
 DB 175 QPNSTYRVAVSVLTPLIOHOMLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS 234  
 QY 339 SRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGQ--PENNYKTTTPVLDSGSFFLYSKLT 396  
 DB 235 SRDELTKSVSVTLINFPADIVHMANRNVSVSEKYNTPPIDADDSYLYSKLT 294  
 QY 397 VDKSRWQOGNVFSCSVNHEALHNHYTQKSLSLSPG 431  
 DB 295 VDKSAMQGTIVTCSVNHEALHNHYTQKSISLTPG 329

## RESULT 20

Ig gamma 1 chain constant region - pig (fragment)  
 147158  
 C:Species: Sus scrofa domestica (domestic pig)  
 C>Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
 C:Accession: I47158  
 R:Kacskovics, I.; Sun, J.; Butler, J.E.  
 J. Immunol. 153, 3565-3573, 1994  
 A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
 A:Reference number: I47158; MUID:95015845; PMID:7930579  
 A:Accession: I47158  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <KAC>  
 A:Cross-references: EMBL:U03778; NID:g433121; PIDN:AAA52216.1; PID:g433122  
 C:Genetics:  
 A:Gene: IGCI  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F.133-202/Domain: immunoglobulin homology <IMV>

Query Match 32.5%; Score 879.5; DB 2; Length 328;  
 Best Local Similarity 55.2%; Pred. No. 3.3e-45;  
 Matches 180; Conservative 38; Mismatches 67; Indels 41; Gaps 7;

```

QY 147 PGSPSPVQGRSPRGKNIQGGKTKT-----LSVQLELDQSGTWTG-----TYLQNK 191
DB 2 PRTAPSVYPLAPCGRDVSGSPNALGCLASYPPEPVYTNWNGALTSGVHTPPSVLQPSG 61
QY 192 KYEFKIDI-VPCPAPEPKS--CDKTY-----TCPELIG-----GPSVL 227
DB 62 LVSLSSMTVVPVPSLSSSKSYTCNVNHPATTKVDKRVGHNQPTGICGCGEAVGSAVFI 121
QY 228 PPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNMYDGEVYNAKTKPREQYNSTRVY 287
DB 122 PPPPKKDTLMISQTPETTCVVDVSKHAEVQFSWYVDGEVYTAETRPKEQFNSTRVY 181
QY 288 VSVLTVLHODMNGKRYKKVSKALPAPLEKTISSAKQOPREPOVYTLPPSRDELTKNQ 347
DB 182 VSVLPTQHDWMLGKFKKKNVNDLPAPITRTISAIQSRPPQVYTLPPRAEELSRK 241
QY 348 VSLTCLVKGFPYSDIAVESNNGQ--PENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQG 405
DB 242 VTLTCLVIGFPPDIVKESKNGQPPENTYRTTPQOVDTGFLYSKLAADKARNMDG 301
QY 406 NVFSCSVMEALHNHYTKSLSPG 431
DB 302 DKFECAVMHEALHNHYTKSLSPG 327

```

## RESULT 21

IG gamma-2b chain C region, membrane-bound form - mouse  
 G2WSBM  
 C:Species: Mus musculus (house mouse)  
 C>Date: 17-Dec-1982 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999  
 C:Accession: C02154; A02158; B02157  
 R:Yamawaki-Kataoka, Y.; Nakai, S.; Miyata, T.; Honjo, T.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 2623-2627, 1982  
 A:Title: Nucleotide sequences of gene segments encoding membrane domains of immunoglobul  
 A:Reference number: A02154; MUID:82222190; PMID:6283537  
 A:Accession: C02154  
 A:Molecule type: DNA  
 A:Residues: 335-405 <YAM>  
 A:Cross-references: GB:J00462  
 R:Rogers, J.; Choi, E.; Souza, L.; Carter, C.; Word, C.; Kuehl, M.; Eisenberg, D.; Wall,  
 Cell 26, 19-27, 1981  
 A:Title: Gene segments encoding transmembrane carboxyl termini of immunoglobulin gamma C  
 A:Reference number: A02158; MUID:82115295; PMID:6799207  
 A:Accession: A02158  
 A:Molecule type: DNA  
 A:Residues: 335-378 <ROG>  
 A:Note: the translation of the first exon of the membrane-bound segment is given

R:Yamawaki-Kataoka, Y.; Kataoka, T.; Takahashi, N.; Obata, M.; Honjo, T.  
 Nature 283, 786-789, 1980

A:Title: Complete nucleotide sequence of immunoglobulin gamma2b chain gene cloned from r  
 A:Reference number: A02157; MUID:80120716; PMID:6766534  
 A:Comments: a allele  
 A:Accession: B02157  
 A:Molecule type: DNA  
 A:Residues: 1-335, 'K' <YA2>  
 C:Comment: The sequence of residues 1-334 was assumed to be identical with the correspo  
 C:Comment: Cell lines producing IgG contain two mRNA species for Ig gamma chains. The m  
 A:introns: 1/1; 98/1; 120/1; 230/1; 335/1; 378/3  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kai  
 C:KeyWords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob  
 F.143-212/Domain: immunoglobulin homology <IMV>  
 F.352-369/Domain: transmembrane #status predicted <TM>  
 F.370-405/Domain: intracellular #status predicted <INT>  
 F.186/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 32.5%; Score 877.5; DB 1; Length 405;  
 Best Local Similarity 51.7%; Pred. No. 5.5e-45;  
 Matches 178; Conservative 52; Mismatches 75; Indels 39; Gaps 9;

```

QY 133 LIQG---QSLTTLSPGSS-----PSVQGRSPRGKNIQGGKTVSVQLELDQSGTW-- 182
DB 28 LVKGYFPSSVYTNWNGSLSSSVHTFPAL-----LQSG-LYTWSSTVTPSPSTPS 77
QY 183 ---TCTVLQ--NQKVEFKID-----IVPCAPPEPKSCDKTHC--PELIGSPVFLF 228
DB 78 QYTVGSVAHAPSSTVDKLEPSGPISTINP-----PCKECHCNPANLEGSPVFI 132
QY 229 PPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYDGEVYNAKTKPREQYNSTRVY 288
DB 133 PENIKDVLMSLTPEVTCVVDVSHEDDPOVQSMFVNVVEVHTAOTHTREDYNSTRIVY 192
QY 289 SVTLTVLHODMNGKRYKKVSKALPAPLEKTISSAKQOPREPOVYTLPPSRDELTKNQ 348
DB 193 STLPTQHDWMLGKFKKKNVNDLPAPITRTISAIQSRPPQVYTLPPRAEQLSRKV 252
QY 349 SLTCLVKGFPYSDIAVESNNGQPPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGNV 408
DB 253 SLTCLVGNPBDIVENTSNQHTENYKDTAPVLDSDGSYITYSKLNKMKTSKMEKTSDF 312
QY 409 SCSVMHEALHNHYTKSLSPGLDQTCABQDGEIDGLWTT 452
DB 313 SCNVNHEGLKNYLLKKTISRSPGLDLDDICABAKDGEIDGLWTT 356

```

## RESULT 22

RMC214  
 T-cell surface glycoprotein CD4 - chimpanzee  
 N:Alternate names: T-cell surface antigen T4/Leu 3  
 C:Species: Pan troglodytes (chimpanzee)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
 C:Accession: B32722; A46534  
 R:Cammerlind, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A:Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the v  
 A:Reference number: A32722; MUID:90182664; PMID:2107024  
 A:Accession: B32722  
 A:Molecule type: mRNA  
 A:Residues: 1-432 <CAM>  
 A:Cross-references: GB:M31135  
 R:Pomsgaard, A.; Hirsch, V.M.; Johnson, P.R.  
 Eur. J. Immunol. 22, 2973-2981, 1992  
 A:Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular rece  
 A:Reference number: A46534; MUID:93049640; PMID:1425921  
 A:Accession: A46534  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA

A:Residues: 3-399 <F0M>  
 A:Note: Sequence extracted from NCBI backbone (NCBI:118332)  
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F:1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAY>  
 F:1-371/Domain: extracellular #status predicted <EXT>  
 F:9-86/Domain: immunoglobulin homology <IM1>  
 F:111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F:191-274/Domain: immunoglobulin homology <IM3>  
 F:296-347/Domain: immunoglobulin homology <IM4>  
 F:372-399/Domain: transmembrane #status predicted <TM>  
 F:396-432/Domain: intracellular #status predicted <INT>  
 F:16-84,130-159,303-345/Dissulfide bonds: #status predicted  
 F:271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.4%; Score 875; DB 1; Length 432;  
 Best Local Similarity 97.7%; Pred. No. 8,4e-45;  
 Matches 170; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 27 KVLGKGDVTELTCTASOKKSIOFHWKNSQIKILNQSGFLTKGPKLMDRADSRSL 86  
 Db 2 KVLGKGDVTELTCTASOKKSIOFHWKNSQIKILNQSGFLTKGPKLMDRADSRSL 61

Qy 87 WDOGNFLLIKNLKIEDSDTYICEVEDQKEEVQLVFGILTANSSTHLLQSGSLTLTLESP 146  
 Db 62 WDOGNFLLIKNLKIEDSDTYICEVEDQKEEVQLVFGILTANSSTHLLQSGSLTLTLESP 121

Qy 147 PGSSSVQCRSPRGKNIQGGKTLVSQLELDSGWTCTVVLONQKVEFKIDIV 200  
 Db 122 PGSSSVQCRSPRGKNIQGGKTLVSQLELDSGWTCTVVLONQKVEFKIDIV 175

RESULT 23

Ig gamma 3 chain constant region - pig (fragment)  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
 C:Accession: I47161  
 R:Kacskovics, I.; Sun, J.; Butler, J.E.  
 J.Immunol. 153, 3565-3573, 1994  
 A>Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
 A:Reference number: I47158; MUID:95015845; PMID:7930579  
 A:Accession: I47161  
 A:Status: preliminary: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-328 <KAC>  
 A:Cross-References: EMBL:U03781; NID:9433127; PIDN:AAA52219.1; PID:9433128  
 C:Genetics:  
 A:Gene: IG33  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:133-202/Domain: immunoglobulin homology <IM>

Query Match 32.3%; Score 873; DB 2; Length 328;  
 Best Local Similarity 57.6%; Pred. No. 8e-45;  
 Matches 179; Conservative 37; Mismatches 61; Indels 34; Gaps 9;

Qy 137 QSLTIT-----LESPPSSPVQCSPPG-KNIQGGKTLVSQLELDSGWTCTVVLONQ 190  
 Db 35 EPLVITMNSGALTSQVHTFPV--LQPSGLYSLSMVTVPASSL--SSKSYTCNV--NH 87

Qy 191 KKEVERID-----IVCPAPEPKSCDKTHTCPCL-LGSPVFLFPKPKDTLMISRT 242  
 Db 88 PATTTKVKRQVGTKTKEPCP-----ICPCCEVAGSIVTFPPKPKDTLMISRT 136

Qy 243 EYTCVVAVDSHEDEPVKFNWYDGYEVNNAKTKPEEYNSYTVVSVLYTLVLAQDMLNGK 302  
 Db 137 EYTCVVAVDSHEDEPVKFNWYDGYEVNNAKTKPEEYNSYTVVSVLYTLVLAQDMLNGK 196

Qy 303 EYKCKVSKALPAPLEKTIKRAKGRPRPOVYTLPPSDELTKNOVSLTCLVKGFPYEDI 362  
 Db 197 EFKCKVNVNVDLPAPITRTISKAIQGRPRPOVYTLPPSDELTKNOVSLTCLVKGFPYEDI 256

Qy 363 AVEWESNGQ--PENNYKTPPVLDSDGSFFLYSKLTVDKSRMOGNVFGSCVMEHALNH 420  
 Db 257 HVEKSKSGQEPESGNVYTTTPQDVDTFFLYSKLAVDKARMDGFECAVMEHALNH 316

Qy 421 YTKQSLSLSPG 431  
 Db 317 YTKQSLSLSPG 327

RESULT 24

G3MSC  
 Ig gamma-3 chain C region, secreted form - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Mar-1987 #sequence\_revision 31-Mar-1991 #text\_change 16-Jul-1999  
 C:Accession: B02156  
 R:Wels, J.A.; Word, C.J.; Rimm, D.; Der-Balan, G.P.; Martinez, H.M.; Tucker, P.W.; Blatch  
 EMBO J. 3, 2041-2046, 1984  
 A>Title: Structure analysis of the murine IgG3 constant region gene.  
 A:Reference number: A02156; MUID:85027161; PMID:6092053  
 A:Accession: B02156  
 A:Molecule type: DNA  
 A:Residues: 1-329 <MEL>  
 A:Cross-References: GB:U00451  
 A:Note: The sequence was determined from the germline gene  
 C:Genetics:  
 A:Insertions: 97/1; 113/1; 223/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)  
 chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu  
 F:19-83/Domain: immunoglobulin homology <IM1>  
 F:97-112/Region: hinge  
 F:136-205/Domain: immunoglobulin homology <IM2>  
 F:247-309/Domain: immunoglobulin homology <IM3>  
 F:179,322/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.7%; Score 857.5; DB 1; Length 329;  
 Best Local Similarity 59.6%; Pred. No. 6,7e-44;  
 Matches 162; Conservative 38; Mismatches 57; Indels 15; Gaps 4;

Qy 170 SVSGLQLDSGWT-----TCTVLONQKVEFKIDI--VCPAPEPKSCDKTHTCP--EL 219  
 Db 62 SLSLVTVVPSSTWPSQTVICNVAPASKTELIRIKRIPKSPSPGS-----SCPGENI 116

Qy 220 LGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSDHEDEVKFNWYDGYEVNNAKTKREE 279  
 Db 117 LGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSDHEDEVKFNWYDGYEVNNAKTKREE 176

Qy 280 QYNSTYRVSVLTVLQDMLNGKKEKCKVSKNKLPAPIEKTISKAKGQPREPQVYTLPPS 339  
 Db 177 QYNSTYRVSVLTVLQDMLNGKKEKCKVSKNKLPAPIEKTISKAKGQPREPQVYTLPPS 236

Qy 340 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDK 399  
 Db 237 REQMSKSKRVSLTCLVNTFFSEALISVEWERNGELEDQYKNTPIILDSGTYFLYSKLTVDK 296

Qy 400 SRMQGQNVFSCVMEHALNHHTQKSLSLSPG 431  
 Db 297 DSWLQGEIFTCVSVMEHALNHHTQKSLSLSPG 328

RESULT 25

PC4436  
 monoclonal antibody 13-1 heavy chain - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 04-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jan-2000  
 C:Accession: PC4436  
 R:Akashi, S.; Kato, K.; Torizama, T.; Dohmae, N.; Yamaguchi, H.; Kamachi, M.; Hazada, A.  
 Biochem. Biophys. Res. Commun. 240, 566-572, 1997  
 A>Title: Structural characterization of mouse monoclonal antibody 13-1 against a porphyr  
 A:Reference number: JCS810; MUID:98063277; PMID:9398605  
 A:Accession: PC4436  
 A:Molecule type: protein

A:Residues: 1-444 <AKA>  
 C:Comment: This catalytic antibody has peroxidase oxidase activity. It is directed against  
 C:Superfamily: Immunoglobulin C region; immunoglobulin homology  
 F:251-320/Domain: immunoglobulin homology <IMM>  
 F:22/Denulfide bonds: Interchain (to 98) #status predicted  
 F:99/Denulfide bonds: Interchain (to 109) #status predicted

Query Match 31.6%; Score 854.5; DB 2; Length 444;  
 Best Local Similarity 41.5%; Pred. No. 1.4e-43;  
 Matches 188; Conservative 76; Mismatches 118; Indels 71; Gaps 14;

QY 30 LGKGDVTELTCTAS--QKKSIOFHW-----KNSQIKL-----GNQGSFLTKGPS 74  
 DB 11 LVRPGNSLKLKSLCTSGFTFSNYRHMILRQPGKLEMIATVTKSDNYGAKVLESVRGRF 70  
 QY 75 KUNDRADSRSLMDQGNFPIIKNLKIEDSDTYIC-----EVED 113  
 DB 71 TIS-RDSSKSSVYLQNM-----RLREEDPATYCCRTPMVYAMDCWQGTIVASSART 123  
 QY 114 QKEEVOLLVFGLTANSDTLLQ-----QSLTTLLESPGSSPSVQCRSPRKNIOG 165  
 DB 124 TPSPVPLPARGSAQNSMTLGLCVKGYPEPVTTMNS--GSLSSGVHTFPA--VLQ 178  
 QY 166 GKTLVSQLELQDSGTW-----TCTVLAQ--NQRKVEFKIDIVCPAPEPKSCDKHTCPE 218  
 DB 179 SDLYTLSSSVTVPSSTWSPSETVTCNVAAHPASSTKVDKKIYPRDQGC-KPCIC---TVEPE 233  
 QY 219 LLAGPSVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEKFMWYVDGVEVNAKTRPRE 278  
 DB 234 V---SSVFIFPPPKVLTITLTPKVCVVDISKDPEVQFSWFDVDEVHTAQTQPRE 290  
 QY 279 EGYNSTYRVSVVTLVHODMLNGEKYCKYKSNKALPAPIEKTISKAKGPREPOVYTLRP 338  
 DB 291 EOPNSTFRSSELPPIHQDMLNGEKYCKRNSAPAPAPIEKTISKAKGPKADQVTTTPR 350  
 QY 339 SRDELTKNOVSLTCLVKGFPYSDIAVWESNGQPENNYKTTPEVLDSGSPFLYSKLTYD 398  
 DB 351 PKQMAKDKVSLTCTMTDFPEDITVEMQNGQPAENYKNTQITMDGSIYFYSLTANQ 410  
 QY 399 KSRWQGNVPSGCVMEHALNHTYOKSLSPG 431  
 DB 411 KSNWEAGNTFTCSVLHEGLNHTKSLSPG 443

## RESULT 26

S31459  
 I9 gamma-1 chain - sheep (fragment)  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
 C:Accession: S31459  
 R:Patri, S.; Nau, F.  
 Submitted to the EMBL Data Library, December 1992  
 A:Reference number: S31459  
 A:Accession: S31459  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-472 <PAT>  
 A:Cross-references: EMBL:X69797  
 C:Superfamily: Immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:277-346/Domain: immunoglobulin homology <IMM>

Query Match 31.6%; Score 854.5; DB 2; Length 472;  
 Best Local Similarity 42.2%; Pred. No. 1.5e-43;  
 Matches 196; Conservative 55; Mismatches 119; Indels 94; Gaps 16;

QY 36 TVELTCTAS--QKKSIOFHWKN---SNQIKILNQG--SFLTKGP---SKUNDRASRS 85  
 DB 34 TISVLTITISGSLANNVGVMDVVRQAPGKALEMLQSGSDIEDIDNPKLSRLSTTKDTSS 93  
 QY 86 LMDQGNFPIIKNLKIEDSDTYICEVEDQKE-----VOLLVFGLTANSDTLL 133  
 DB 94 -----QVSLTLSTVTTEDTAIVVYCARVDSDSHAFAYASYDFWGPGLLISVLSAST----- 144

QY 134 LQGSUTLLESPPGSSPVSQCRSPRKNI-----QSGKTLVSQ 173  
 DB 145 -----TPPKVYPLTSCCGDSSISVTLGLCVSSYMEPEVVTMNSALTSIGVHT 193

QY 174 LE--LQDSG-----TWCTVLAQ--NQRKVEFKIDIVCPAPEPKSCD 211  
 DB 194 PFAILLQSGSLVLSVTVTPASTSGAQFTICNVAAHPASSTKVDKRIE--GCPDP--CKHC- 250

QY 212 KTHTC--PELLGSPVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEKFMWYVDGVE 269  
 DB 251 ---RCPPPLPGSPSVIFPPPKKDTLITSGPEVTCVVDVGQDDPEVQFSWFDVNV 307

QY 270 HNAKTRPREQYNSTYRVSVVTLVHODMLNGEKYCKYKSNKALPAPIEKTISKAKGPR 329  
 DB 308 FRARKTRPREQYNSTYRVSVVTLVHODMLNGEKYCKYKSNKALPAPIEKTISKAKGPR 367

QY 330 EPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVWESNGQPE--ENNYKTTPEVLDSG 387  
 DB 368 EPQVYVLAIPQELSKSLSTVTCVLTGFYPDVIAVEMQNGQPESEDKYGTTTSLQDADG 427

QY 388 SFFLYSKLTVDSKRWQGNVPSGCVMEHALNHTYOKSLSPG 431  
 DB 428 SYFLYSRLRVDRKNSWQEGDPTVACVWMEHALNHTYOKSLSPG 471

## RESULT 27

S37483  
 I9 gamma-2a chain - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
 C:Accession: S37483  
 R:Pucancel, F.F.D.  
 Submitted to the EMBL Data Library, February 1993  
 A:Reference number: S37483  
 A:Accession: S37483  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-469 <DUC>  
 A:Cross-references: EMBL:X70423; NID:g406252; PIDN:CAA9868.1; PID:g406253  
 C:Superfamily: Immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:276-345/Domain: immunoglobulin homology <IMM>

Query Match 31.5%; Score 850.5; DB 2; Length 469;  
 Best Local Similarity 42.5%; Pred. No. 2.7e-43;  
 Matches 197; Conservative 56; Mismatches 125; Indels 85; Gaps 15;

QY 30 LGKGDVTELTCTASQKKSIOFHWKNSQIKILNQG-----SFLTKGPSLMDRADS 83  
 DB 30 LVRPGASVKISCKASGTYFTDY---INNVKQKPGQGLKIMWIYPASGNTYNNFKKG 86  
 QY 84 RSL---WDQGNFPIIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDTLLQ--GQS 138  
 DB 87 ATLVTDTSSSTAYMQLSLSIEDTAIFC-----ARMGATAT---LIDWQGG 132  
 QY 139 LTLLESPPGSSPSVQ-----CRSPRKNIOG-----KTL----- 169  
 DB 133 TLTIVSSAKTAPASVPLAPVCGDTTSSVTLGLCVKGYPEPVTLTWNSSLSGVHTF 192  
 QY 170 -SVSQLEQ-----DSGTW-----TCTVLAQ--NQRKVEFKIDIVCPAPEPK 208  
 DB 193 PAVLQSDLTLSSTVTVTSTWSPSITCNVAHPASSTVDRKRIEPRGFTIKPCP--PC 249  
 QY 209 SCDKTHTCPELLGSPVFLFPPPKKDTLMISRTPEVTCVVDVSHEDPEKFMWYVDGVE 268  
 DB 250 KC-----PAPNLGGSVFIFFPKKIDVLMISLPIVTCVVDVSDDDPVQVSWVNNV 305  
 QY 269 VNAKTRPREQYNSTYRVSVVTLVHODMLNGEKYCKYKSNKALPAPIEKTISKAKGPR 328  
 DB 306 VHTAQTOTHTREDYNSLTIRVVSALPIQHDPMNSGKEFKCYNNKDLPAPIERTISKPKGSV 365  
 QY 329 REPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVWESNGQPENNYKTTPEVLDSGS 388

Db 366 RAPQYVLPPEEETKQVLTICVTFMPEPDIYVETNNKGTETLANKTEPVLDSGS 425  
 QY 369 PFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSPG 431  
 Db 426 YFMSKLRKVKRKNWERNUSYSCSVHEGHLNHYTOKSLSPG 468

## RESULT 28

C30554

Ig heavy chain C region - sheep (fragment)

C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C/Date: 03-Mar-1989 #sequence\_revision 03-Mar-1989 #text\_change 21-Jan-2000

C/Accession: C30554

R/Foley, R.C.; Beh, K.J.

J. Immunol. 142, 708-711, 1989

A/Title: Isolation and sequence of sheep Ig H and L chain cDNA.

A/Reference number: A30554; MUID:89093962; PMID:2492052

A/Accession: C30554

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-308 &lt;FOL&gt;

C/Superfamily: immunoglobulin C region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F/113-182/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 31.0%; Score 838; DB 2; Length 308;  
 Best Local Similarity 60.7%; Pred. No. 8,9e-43;  
 Matches 164; Conservative 37; Mismatches 55; Indels 14; Gaps 7;

QY 170 SVSLELDGSG--TWCTVLQ--NOKYEFKIDIVCPAPEKSCDKHTHC--PELLGAP 223  
 Db 44 SVTVAPASTSGAQTFCVNAHPASTKVDKRYE--GCGDP--CKHC----RCPPPELPGP 97  
 QY 224 SVFLFPPKPKDTLMSRTPEVTCVVDVSHEDPEKFMVYDGVVNAKTPREEOYNS 283  
 Db 98 SVFIFPPKPKDTLTISGPEVTCVVDVQDDPEVQFSPVNDVNEVTAATKPREEOYNS 157  
 QY 284 TYRVSVLTVLHODMNGEKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDEL 343  
 Db 158 TFRVVSALPIQHODMNGEKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDEL 217  
 QY 344 TKNOVSLTCLVKGFPSPDIAMWESNGQP--ENNYKTTTPVLDSGSEFLYSKLTVDKSR 401  
 Db 218 SKSTLSVTCVLTGFPDYIAVEMQNGQDESEDKYGTTSQDADGSEFLYSRLKVDNS 277  
 QY 402 WQGNVFCSCVMHEALHNHYTOKSLSPG 431  
 Db 278 WQEGDTYACVMHEALHNHYTOKSLSPG 307

## RESULT 29

PS0018

Ig gamma-2b chain C region - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 16-Jul-1999

C/Accession: PS0018; B25941

R/Brueggemann, M.

Gene 74, 473-482, 1988

A/Title: Evolution of the rat immunoglobulin gamma heavy-chain gene family.

A/Reference number: PS0017; MUID:89232738; PMID:3149946

A/Accession: PS0018

A/Molecule type: DNA

A/Residues: 1-333 &lt;BRU&gt;

R/Brueggemann, M.; Free, J.; Diamond, A.; Howard, J.; Cobbold, S.; Waldmann, H.

Proc. Natl. Acad. Sci. U.S.A. 85, 6075-6079, 1988

A/Title: Immunoglobulin heavy chain locus of the rat: striking homology to mouse antibody

A/Reference number: A25941; MUID:86287297; PMID:3016742

A/Accession: B25941

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 227-333 &lt;BR2&gt;

C/Genetics:

A/Introns: 96/1; 117/1; 227/1  
 C/Superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: immunoglobulin  
 F/20-82/Domain: immunoglobulin homology <IMM>

Query Match 30.8%; Score 833; DB 2; Length 333;  
 Best Local Similarity 58.6%; Pred. No. 1,9e-42;  
 Matches 157; Conservative 38; Mismatches 53; Indels 20; Gaps 5;

QY 179 SGTW-----TCTVQ--NOKYEFKID-----VPCPAPPEKSCDKHTHC--PELLGAP 223  
 Db 70 SFTVPSQTVTCVNAHPASTKVDKRYERNGGIGHKCP----TCTPCHKCPVELLGGP 124  
 QY 224 SVFLFPPKPKDTLMSRTPEVTCVVDVSHEDPEKFMVYDGVVNAKTPREEOYNS 283  
 Db 125 SVFIFPPKPKDTLTISGNNAVTCVVDVSEEPVQFSPVNDVNEVTAATKPREEOYNS 184  
 QY 284 TYRVSVLTVLHODMNGEKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDEL 343  
 Db 185 TFRVVSALPIQHODMNGEKYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDEL 244  
 QY 344 TKNOVSLTCLVKGFPSPDIAMWESNGQPRENNYKTTTPVLDSGSEFLYSKLTVDKSRWQ 403  
 Db 245 TEQTVSLTCLVKGFPSPDIAMWESNGQPRENNYKTTTPVLDSGSEFLYSKLTVDKSRWQ 304  
 QY 404 QGNVFCSCVMHEALHNHYTOKSLSPG 431  
 Db 305 SRAVFCSCVMHEALHNHYTOKSLSPG 332

## RESULT 30

S40295

Ig gamma-2a chain (mAb735) - mouse

C/Species: Mus musculus (house mouse)

C/Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 16-Jul-1999

C/Accession: S40295

R/Kiebert, S.; Kraztzn, H.D.; Zimmermann, B.; Vaesen, M.; Frosch, M.; Weisgerber, C.; Bl

submitted to the EMBL Data Library, January 1993

A/Description: Primary structure of the murine monoclonal IgG2a antibody mAb735 against

A/Reference number: S40295

A/Accession: S40295

A/Molecule type: protein

A/Residues: 1-446 &lt;KLE&gt;

C/Genetics:

A/Map position: 12

C/Superfamily: immunoglobulin C region; immunoglobulin homology

C/Keywords: disulfide bond; glycoprotein; immunoglobulin; pyroglutamic acid

F/1-446/Product: Ig gamma-2a chain #status experimental &lt;MAT&gt;

F/1-117/Domain: V-D-J region &lt;VDJ&gt;

F/118-446/Domain: C region &lt;CHR&gt;

F/118-214/Domain: C1 region &lt;CH1&gt;

F/215-230/Region: hinge

F/231-340/Domain: C2 region &lt;CH2&gt;

F/341-446/Domain: C3 region &lt;CH3&gt;

F/360-427/Domain: immunoglobulin homology &lt;IMM&gt;

F/1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

F/222-96,144-199,261-321,367-428/Disulfide bonds: #status predicted

F/113/Disulfide bonds: interchain (to light chain) #status predicted

F/224,227,229/Disulfide bonds: interchain #status predicted

F/297/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 30.5%; Score 825; DB 2; Length 446;  
 Best Local Similarity 40.9%; Pred. No. 8,2e-42;  
 Matches 190; Conservative 60; Mismatches 123; Indels 92; Gaps 15;

QY 30 LGKKGDTVELTCTASQKSIQF--HMKNQIKILGNQ-----SFLTKGPSKLNDRAD 81  
 Db 11 LVRRGASVSKSKSGVTFDYIHW-----VKQRPBGELWIGMIVPGSNTVKNKEXFK 65  
 QY 82 SRRL---WQGNPPLIKUKIEDSPYICEVDDQKEVQVLVFGLTANSDTHLLO--G 136  
 Db 66 GKATLTVDTSSTAYWQSLSTSEDSAVYFC-----ARGSKPAMDYWG 108



```

Qy 137 QSLTTLSPSSPPSSPVQ-----CRSPRGKNTIQG-----KTL----- 169
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 109 QGRTSVSSAKTAPSVYPLAPVCGDTTSSVTLGLGVGYPEPVTLTMTNSSLSSGVH 168
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 170 ---SVSQLELQ-----DSGTW-----TCTVLO--NOKKVEFKID-----IVPCDAPB 206
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 TEPVALQSDLYTLSSSVTSTSTWPSQSITCNVAHPASSTKVDKKEPRGPTTKPCP--- 225
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 207 PKSCDKTHNCPCLLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNNYV 266
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 226 PCKC-----PAPNLLGSPSVFIFPPPKIKDVLMTSLSMVTCVVVDVSEDDPDVQISFVN 281
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 267 VEVNNAKTREREOQNSTYRVVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISKAG 326
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 282 VEVLTAQTOTHRDYNSTLRVVSALPFIQODWMSGKEPCKVNNKOLPAPIETKISKPG 341
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 327 QRPPEPVYTLPPSRDELTKQVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPYLSD 386
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 342 SVRAAPQVYVLPPEEEMTKQVTLTGMVTD FMEPEDIYVEWTNNGKTELNYKNTPEVLDSD 401
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 387 GSPFLYSKLTVDKSRWQGVPSGVNHEALHNHYTOKSLSPG 431
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 402 GSTFMTSKLRVEKQKNVERNSYCSVHGLNHNHTTKSFSPRTG 446
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 31
PS0017
Ig gamma-1 chain C region - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 16-Jul-1999
C/Accession: PS0017; C25941
R.Brueggemann, M.
Gene 74, 473-482, 1988
A>Title: Evolution of the rat immunoglobulin gamma heavy-chain gene family.
A/Reference number: PS0017; MUID:89232738; PMID:3149946
A/Accession: PS0017
A/Molecule type: DNA
A/Residues: 1-326 <BRU>
R.Brueggemann, M.; Frey, J.; Diamond, A.; Howard, J.; Cobbold, S.; Waldmann, H.
Proc. Natl. Acad. Sci. U.S.A. 83, 6075-6079, 1986
A>Title: Immunoglobulin heavy chain locus of the rat: striking homology to mouse antibody
A/Reference number: A25941; MUID:86287397; PMID:3016742
A/Accession: C25941
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 220-326 <BR2>
C/Genetics:
A/Introns: 98/1; 113/1; 220/1
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin
F/20-84/Domain: immunoglobulin homology <IMM>

Query Match 30.5%; Score 823.5; DB 2; Length 326;
Best Local Similarity 50.0%; Pred. No. 6, 9e-42;
Matches 163; Conservative 53; Mismatches 55; Indels 55; Gaps 9;

Qy 133 LLOG---GSLTLTLEPPSS-----PSVQCSPPKAKTIQGGKTLSSVQLELDSGTW-- 182
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 LVKGFPEPEVTVNMGALSSGVNTPPAV-----LQSGLYTLTSSVTV-PSSTWPS 77
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 183 ---TCTVLO--NOKKVEFKIDIVPCAPAPKSCDKTHNCPCLLGG-----PSV 225
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 QVTTCVNAHPASSTKVDKTI-----VPRNC-----GGCKPCTCTGSEVSSV 119
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 226 FLFPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNNYVGVENVNAKTREREOQNSTY 285
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 FIFPPKPKDVLTLTPKTCVVDVSDDPVHFGMFDVENVHQAQTRPREQGNSTF 179
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 286 RVVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISKAGQRPPEQVYTLPPSRDELTK 345
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 RSVSELPILHODWLNKTRCKVTSAPSPLEKITSKEGRTQVPHVYVMSPTKEEMTQ 239
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 346 NQVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPYLSDSGSFLYSKLTVDKSRWQGV 405

```

```

Db 240 NEVSITCWKGFYPPDIYVEMQNGPQENYKNTPTMTDTSFYLYSLNAYKEKMQG 299
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 406 NVPSCSVNHEALHNHYTOKSLSPG 431
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 NTFCSVLHGLNHNHTKSLSHSPG 325
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 32
S00847
Ig gamma-2c chain C region - rat (fragment)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 23-Jul-1999
C/Accession: S00847
R.Brueggemann, M.; Delmasro-Galfre, P.; Waldmann, H.; Calabi, F.
Eur. J. Immunol. 18, 317-319, 1988
A>Title: Sequence of a rat immunoglobulin gamma-2c heavy chain constant region cDNA: ex
A/Reference number: S00847; MUID:88166903; PMID:3127222
A/Accession: S00847
A/Molecule type: mRNA
A/Residues: 1-329 <BRU>
A/Cross-references: EMBL:X07189; NID:G57602; PID:CAA30169.1; PID:G663228
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin
F/20-84/Domain: immunoglobulin homology <IMM>

Query Match 30.5%; Score 823; DB 2; Length 329;
Best Local Similarity 58.5%; Pred. No. 7, 5e-42;
Matches 151; Conservative 45; Mismatches 56; Indels 6; Gaps 2;

Qy 179 SGTW-----TCTVLOQKVEFKIDIVPCAPAPKSCDKTHNCPCLLGGPSVFLPFPKPK 233
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 SSTWSSQVTCVNAHPATSNLKRIEP-RRKRPRTDSCDDVLRPSVFIPEPKK 130
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 234 DTLMTSRPEVTCVVVDVSHEDPEVKFNNYVDGVEVNAKTREREOQNSTYRVVSVLT 293
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 DILMTITLTPKTCVVDVSEEDPDVQFSWFDVNVFTAQOTQPHBEQLNGTFRVYSTLHI 190
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 294 LHODWLNKGEYCKVSNKALPAPIETKISKAGQRPPEQVYTLPPSRDELTKQVSLTCL 353
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 QHODWMSGKEPCKVNNKOLPAPIETKISKPRGKATPVYTIPPREQMSKXVSLTGM 250
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 354 VKGFYPSDIAVWESNGQPENNYKTPPYLSDSGSFLYSKLTVDKSRWQGVPSGVN 413
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 VTSFTPASISVEMERNGELDQYKNTLPVLDSDSESYFLYSKLSVDTDSMMRDYITCSV 310
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 414 HEALHNHYTOKSLSPG 431
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 HEALHNHNHTKSLSPG 328
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 33
GIMS
Ig gamma-1 chain C region, secreted form - mouse
C/Species: Mus musculus (house mouse)
C/Date: 31-Mar-1980 #sequence_revision 24-Sep-1981 #text_change 16-Jul-1999
C/Accession: A02159; A26234; A26236
R.Honjo, T.; Obata, M.; Yamawaki-Kataoka, Y.; Kataoka, T.; Kawakami, T.; Takahashi, N.;
Cell 18, 559-568, 1979
A>Title: Cloning and complete nucleotide sequence of mouse immunoglobulin gamma1 chain c
A/Reference number: A02159; MUID:80045036; PMID:115553
A/Accession: A02159
A/Molecule type: DNA
A/Residues: 1-324 <HON>
A/Cross-references: GB:J00453
A/Note: the sequence was determined from the germ-line gene
A/Note: Lys-324 is removed posttranslationally
R.Obata, M.; Yamawaki-Kataoka, Y.; Takahashi, N.; Kataoka, T.; Shimizu, A.; Mano, Y.; Se
Gene 9, 87-97, 1980
A>Title: Immunoglobulin gamma-1 heavy chain gene: structural gene sequences cloned in a
A/Reference number: A26234; MUID:80202559; PMID:6769752
A/Contents: MOPC 31C
A/Accession: A26234

```

```

A:Molecule type: mRNA
A:Residues: 76-324 <OBA>
A:Cross-references: GB:V00775; NID:951652; PIDN:CAA24153.1; PID:951653
R:Roger, J.; Clarke, P.; Salsner, W.
Nucleic Acids Res. 6, 3305-3321, 1979
A:Title: Sequence analysis of cloned cDNA encoding part of an immunoglobulin heavy chain
A:Reference number: A26236; MUID:80012837; PMID:113776
A:Contents: MOPC 21
A:Accession: A26236
A:Molecule type: mRNA
A:Residues: 170-275, 'D', 277, 'D', 279-322 <ROG>
A:Cross-references: GB:V00795; NID:951830; PIDN:CAA24176.1; PID:9780265
R:Adetugbo, K.
J. Biol. Chem. 253, 6068-6075, 1978
A:Title: Evolution of immunoglobulin subclasses. Primary structure of a murine myeloma g
A:Reference number: A26237; MUID:78242288; PMID:98524
A:Contents: annotation; MOPC 21
A>Note: This is the final paper in a series reporting the protein sequence, the disulfid
A>Note: There are a number of differences from the sequence shown
C:Genetics:
A:Introns: 1/1; 98/1; 111/1; 218/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 1a
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob
F:20-84/Domain: immunoglobulin homology <IM1>
F:98-110/Region: hinge
F:131-200/Domain: immunoglobulin homology <IM2>
F:237-304/Domain: immunoglobulin homology <IM3>
F:27-82, 138-198, 244-302/Disulfide bonds: #status experimental
F:102/Disulfide bonds: interchain (to light chain) #status experimental
F:104, 107, 109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:174/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query March 30.3%; Score 818.5; DB 1; Length 324;
Best Local Similarity 51.8%; Pred. No. 1.4e-41;
Matches 156; Conservative 50; Mismatches 60; Indels 35; Gaps 5;

QY 133 LLQGQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDSDGTWCTVLAQ--NQ 190
Db 56 VLQGDVLTLSSTVTVVPSPP-----RPSITVCNVAHPASS 90
QY 191 KKEVEKIDIVCPAPEPEKSCDKHTTCPELLGSPVLEPPPKDKTLMISRTPEVTCVVD 250
Db 91 TKVDKKTIVPRDCCG-KPCIC---TVPEV---SSVFIPPKPKDVLITTLTPKTCVVD 142
QY 251 VSHDEPEVKEMVYDGVGVHNAKTKPREOYNSTRVSVLTVAHOMDLNKEKKCKRYSN 310
Db 143 ISKDPPEVQFSFVDDVDEVHTAOTPREEPNSTFRSVSELPIMHODMLNKEKRCRYNS 202
QY 311 KALPAIEKTSKAKGQPREPOVYTLTPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 370
Db 203 AAFAPAIETKISKRGKAPQVYTIIPPKQMAKDKSLCTMTDPEDITYEMQNG 262
QY 371 QPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCGVMEHALNHYTKSLISP 430
Db 263 QPAENYKATQIPIMNTNGSYFYYSKLVNQKSNMEAGNFTCSVLHGLAHNHTKSLSHSP 322
QY 431 G 431
Db 323 G 323

RESULT 34
G2MS11
Ig gamma-2b chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence revision 01-Dec-2000 #text change 01-Dec-2000
C:Accession: S25057; A02157; A26235; A26232; A53598
R:Flischer, R.; Voss, A.; Niersbach, M.; Munkler, W.; Hirsch, H.J.; Kreuzaler, F.
submitted to the EMBL Data Library, July 1992
A:Description: Production of a Tobacco mosaic virus (TMV) inactivating neotop specific m
A:Reference number: S25057

```

```

A:Accession: S25057
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-474 <FIS>
A:Cross-references: EMBL:X67210; NID:954826; PIDN:CAA47649.1; PID:954827
R:Yanawaki, Katoaka, Y.; Katoaka, T.; Takahashi, N.; Obata, M.; Honjo, T.
Nature 283, 788-789, 1980
A:Title: Complete nucleotide sequence of immunoglobulin gamma2b chain cloned from n
A:Reference number: A02157; MUID:80120716; PMID:676534
A:Contents: a allele
A:Accession: A02157
A:Molecule type: DNA
A:Residues: 138-161, 'L', 163-189, 'FP', 193-474 <YAM>
A:Cross-references: GB:U00461
A>Note: The sequences were determined from the germline gene
R:Tucker, P.W.; Marcu, K.B.; Slightom, J.L.; Blattner, F.R.
Science 206, 1299-1303, 1979
A:Title: Structure of the constant and 3' untranslated regions of the murine gamma2b hea
A:Reference number: A26235; MUID:80081501; PMID:117548
A:Contents: MPC 11
A:Accession: A26235
A:Molecule type: mRNA
A:Residues: 138-172, 'P', 174-189, 'FP', 193-376, 'T', 378-474 <TU1>
A>Note: Lys-474 is probably removed posttranslationally
R:Tucker, P.W.; Marcu, K.B.; Newell, N.; Richards, J.; Blattner, F.R.
Science 206, 1303-1306, 1979
A:Title: Sequence of the cloned gene for the constant region of murine gamma2b immunoglob
A:Reference number: A26232; MUID:80081502; PMID:117549
A:Accession: A26232
A:Molecule type: DNA
A:Residues: 138-172, 'P', 174-189, 'FP', 193-376, 'T', 378-474 <TU2>
R:Ollo, R.; Rougeon, F.
Nature 296, 761-763, 1982
A:Title: Mouse immunoglobulin allotypes: post-duplication divergence of gamma2a and gamma
A:Reference number: A26233; MUID:82179203; PMID:6803173
A:Contents: b allele
A:Accession: A26233
A:Molecule type: DNA
A:Residues: 138-161, 'L', 163-189, 'FP', 193-300, 'R', 302-331, 'A', 333-437, 'DI', 440-474 <OL1>
R:Kim, H.; Yamaguchi, Y.; Masuda, K.; Matsunaga, C.; Yamamoto, K.; Irimura, T.; Takahash
J. Biol. Chem. 269, 12345-12350, 1994
A:Title: O-glycosylation in hinge region of mouse immunoglobulin G2b.
A:Reference number: A53598; MUID:94216359; PMID:7512967
A:Accession: A53598
A:Status: preliminary
A:Molecule type: protein
A:Residues: 234-251 <KIM>
C:Comment: The a allele sequence is shown.
C:Genetics:
A:Introns: 138/1; 236/1; 258/1; 368/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 1a
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F:157-222/Domain: immunoglobulin homology <IM1>
F:236-257/Region: hinge
F:281-350/Domain: immunoglobulin homology <IM2>
F:387-454/Domain: immunoglobulin homology <IM3>
F:152/Disulfide bonds: interchain (to light chain) #status predicted
F:154-220, 288-348, 394-452/Disulfide bonds: #status predicted
F:247, 250, 253, 256/Disulfide bonds: interchain (to heavy chain) #status predicted
F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query March 30.3%; Score 818; DB 1; Length 474;
Best Local Similarity 39.7%; Pred. No. 2.3e-41;
Matches 184; Conservative 66; Mismatches 124; Indels 90; Gaps 14;

QY 34 GDTVELCTASQKSIQF--HMKNNOIKLNGG----SFL--TKGPKNDPADSRSS 85
Db 34 GASYKMSCKASGYTFITVYVHW-----VKQKPGGSLGWIGYINNKQGTKEKFKGKAT 88
QY 86 LW--DQGNPLIINKLKIEDSDTYICEVBDQKEEVQLLVGLTRANSDTHLQSGSLTLT 142

```

```

Db      89 LPSDKSNATAYWELSLTSEDSAVVYCCARD-----YDWMFAVWGGGLVT 134
Qy      143 LESPSSSSSVQ-----CSPGKAIQGG-----KTLVSQLE 175
Db      135 VSAAKTTPSVVPLAPGCGDITGSSVTSQCLVKGYPPESVTVTWSGSLSSVHTLSQL 194
Qy      176 LOD-----SGTW-----TCTVQ--NOKKVEKID-----IYCPAPEPS 209
Db      195 LQSGLYTSSSVTVSSSTPSCQTVTCVNHPSSTTVDDKLEPSPGISTINPCP-----P 249
Qy      210 CDKTHTC--PELLGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNYYDGV 267
Db      250 CKECHKCAPNLEGGGSVIFPPNKKDVLMIISLTPTCVVVDVSEDDPDVQISWFNANNV 309
Qy      268 EVHNAKTKREEQYNSTYVSVLVTLVHODMLNGEKYCKKSNKALPAPIEKTISKAKQ 327
Db      310 EVHTAQOTQTHREDYNSTIRVSTLPIQHODMWSGKEFKCKVNNKDLPSPIERTISKIKGL 369
Qy      328 PREPOVYTLPSRDELTKNQSILTLGVGFYPSDIAVEMESNGOPENNYKTPVLDSDG 387
Db      370 VRAPOVYTLPPAEQUSKRDVSLTLVGFNPDGDSVEMTNGHTBEANTKDTAPVLDSDG 429
Qy      388 SFFLYSKLTVDKSRMOQGNVFCSCVWHEALHNHYTQKSLSLSPG 431
Db      430 SYFIYSKLMNMTSKMEKTDSPFCGNVHBEGLKNYILKKTISRSPG 473

```

## RESULT 35

GMSA

```

Ig gamma-2a chain C region, secreted form (allele a) - mouse
C.Species: Mus musculus (house mouse)
C.Date: 30-Sep-1980 #sequence_revision 01-Sep-1981 #text_change 16-Jun-2000
C.Accession: A02152; A32657; A32658
R.Sikorav, J.L.; Aufrey, C.; Rougeon, F.
Nucleic Acids Res. 8, 3143-3155, 1980
A.Title: Structure of the constant and 3' untranslated regions of the murine Balb/C gamma
A.Reference number: A02152; MUID:81076554; PMID:677755
A.Accession: A02152
A.Molecule type: mRNA
A.Residues: 1-330 <SIK>
A.Cross-references: GB:V00798; NID:G51835; PIDN:CNA24178.1; PID:G1333984
R.Ollo, R.; Aufrey, C.; Morchamps, C.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 78, 2442-2446, 1981
A.Title: Comparison of mouse immunoglobulin gamma-2A and gamma-2B chain genes suggests t
A.Reference number: A32657; MUID:81198976; PMID:6262729
A.Accession: A32657
A.Molecule type: DNA
A.Residues: 1-330 <YAM>
A.Cross-references: GB:J00470
A.Note: The sequence was determined from the germ-line gene
R.Ollo, R.; Aufrey, C.; Morchamps, C.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 78, 2442-2446, 1981
A.Title: Comparison of mouse immunoglobulin gamma-2A and gamma-2B chain genes suggests t
A.Reference number: A32658; MUID:81232894; PMID:6787604
A.Accession: A32658
A.Molecule type: DNA
A.Residues: 1-330 <OLJ>
A.Note: The sequence was determined from the germ-line gene
R.Bourgeois, A.; Fougereau, M.; Rocca-Serra, J.
Eur. J. Biochem. 43, 423-435, 1974
A.Title: Determination of the primary structure of a mouse IgG2a immunoglobulin: amino-a
A.Reference number: A32659; MUID:74175517; PMID:4831970
A.Contents: annotation; myeloma protein MOPC 173
A.Note: this is one paper in a series reporting the sequence; for additional references,
Ride Preval, C.; Fougereau, M.
Eur. J. Biochem. 30, 452-462, 1972
A.Title: Determination of the primary structure of a mouse gamma2a immunoglobulin. Iden
A.Reference number: A32660; MUID:73056887; PMID:4565406
A.Contents: annotation; MOPC 173, disulfide bonds
C.Genetics:

```

```

A.Introns: 1/1; 98/1; 114/1; 224/1
C.Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C.Superfamily: immunoglobulin C region; immunoglobulin homology
C.Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F:20-84/Domain: immunoglobulin homology <IM1>
F:98-113/Region: hinge
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:15/Disulfide bonds: interchain (to light chain) #status experimental
F:27-82,144-204,250-308/Disulfide bonds: #status experimental
F:107,116,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match      29.8%; Score 805.5; DB 1; Length 330;
Best Local Similarity 58.5%; Pred. No. 8,2e-41;
Matches 155; Conservative 33; Mismatches 58; Indels 19; Gaps 5;

```

```

Qy      179 SGTW-----TCTVLD--NOKKVEKID-----IYCPAPEPSGDKTHTCPELLGSPSVF 226
Db      72 SGTWPSQSIITCVNHPASTKYDKIEBGRPIKPCP---PCKC---PAPVLDGSPSVF 124
Qy      227 LPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNYYDGVENHNAKTKPREEQYNSTYR 286
Db      125 IFPPKIKDVLMIISLPIVTCVVVDVSEDDPDVQISWFNANNVEHTAQOTHREDYNSTLR 184
Qy      287 VVSUTVTLHODMLNGEKYCKKSNKALPAPIEKTISKAKQPREQVYTLPPSRDELTKN 346
Db      185 VVSALPIQHODMWSGKEFKCKVNNKDLPAPIERTISKRGSVRAPOVYVLPPEPEMTK 244
Qy      347 QVSLTCLVKGFPSPDIAVEMESNGOPENNYKTPVLDSDGSFFLYSKLTVDKSRMOQN 406
Db      245 QVTLTCMVDTFWPEDIIVETWNGTETLNTKTEPVLDSGYSFYMSKLRVKKMVERN 304
Qy      407 VFSCSVWHEALHNHYTQKSLSLSPG 431
Db      305 SVSCSVVHBEGLKNYILKKTISRSPG 329

```

## RESULT 36

GMSAB

```

Ig gamma-2a chain C region, secreted form (allele b) - mouse
C.Species: Mus musculus (house mouse)
C.Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
C.Accession: A02153; A32656
R.Schreier, P.H.; Botwell, A.L.M.; Mueller-Hill, B.; Baltimore, D.
Proc. Natl. Acad. Sci. U.S.A. 78, 4495-4499, 1981
A.Title: Multiple differences between the nucleic acid sequences of the IgG2a(a) and IgC
A.Reference number: A02153; MUID:82037861; PMID:6170065
A.Accession: A02153
A.Molecule type: mRNA
A.Residues: 1-335 <SCH>
A.Cross-references: GB:J00479
A.Experimental source: strain C57BL/6
R.Dognin, M.J.; Laureys, M.; Stroberg, A.D.
Proc. Natl. Acad. Sci. U.S.A. 78, 4031-4035, 1981
A>Title: Multiple amino acid substitutions between murine gamma 2a heavy chain Fc region
A.Reference number: A32656; MUID:82037777; PMID:6794027
A.Accession: A32656
A.Molecule type: protein
A.Residues: 118-267; 'E', 269-328; 'G', 330-334 <DOG>
C.Comment: Lys-335 is removed posttranslationally.
C.Complex: The sequence differs from that of the allele, from BALB/c mice, at 15% of t
C.Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C.Superfamily: immunoglobulin C region; immunoglobulin homology
C.Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F:20-84/Domain: immunoglobulin homology <IM1>
F:98-118/Region: hinge
F:142-211/Domain: immunoglobulin homology <IM2>
F:248-315/Domain: immunoglobulin homology <IM3>
F:15/Disulfide bonds: interchain (to light chain) #status predicted
F:27-82,149-209,255-313/Disulfide bonds: #status predicted

```

F108,117/Disulfide bonds: Interchain (to heavy chain) #status predicted  
F185/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 29.6%; Score 801; DB 1; Length 335;

Best Local Similarity 50.2%; Pred. No. 1.5e-40;

Matches 166; Conservative 52; Mismatches 77; Indels 36; Gaps 9;

```

QY 124 GLTANSDT--HLQO---QSLTLTLESPGSS-----PSVOCSPRGKNIQSGKTLVSQ 173
DB 17 GTTSSVTLGLGVKGFPEPVLTLWNSSGLSSGVHTFPAL-----LQSG-LYTLSS 66
QY 174 LELQDSGTW-----TCTV-----LQNKQVEFKIDY--PCPAPEKSCDKHTCPRL 220
DB 67 SVTVSNMTPSQTITCNVAHPASSTKVAKIEPRVITQNPCC---PQGRVPCAPDL 123
QY 221 GGPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVNAHAKKPREQ 280
DB 124 GGPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYDGVNAHAKKPREQ 183
QY 281 YNSTRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSR 340
DB 184 YNSTRVVSVLPIQHODMNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSR 243
QY 341 DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKS 400
DB 244 EEMTKKEFSLTCMTITGFLPAEIAVDWMTSGRTEQVKYATVLDSDGSFFLYSKLTVDKS 303
QY 401 RMOQGNVFGSCVMHEALHNHYTOKSLSLSPG 431
DB 304 TWERGSFLPACSVHVELNHLNLTTKTISRSLG 334

```

#### RESULT 37

S01321

Ig gamma-2b chain precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 23-Jul-1999

C:Accession: S01321

R:de Waele, P.; Feys, V.; van de Voorde, A.; Molemans, F.; Fiers, W.

Eur. J. Biochem. 176, 287-295, 1988

A:Title: Expression in non-lymphoid cells of mouse recombinant immunoglobulin directed at

A:Reference number: S01320; MUID:88329081; PMID:3138116

A:Accession: S01321

A:Molecule type: mRNA

A:Residues: 1-475 <DE1>

A:Cross-references: EMBL:X13188; NID:G51780; PID:CA31580.1; PID:G51781

A>Note: this sequence was determined from the differentiated gene

C:Superfamily: Immunoglobulin C region; Immunoglobulin homology

C:Keywords: Immunoglobulin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-475/Product: Ig gamma-2b chain #status predicted <MAT>

F:159-223/Domain: immunoglobulin homology <IMM>

Query Match 29.4%; Score 793.5; DB 2; Length 475;

Best Local Similarity 40.0%; Pred. No. 6.5e-40;

Matches 187; Conservative 58; Mismatches 135; Indels 87; Gaps 14;

```

QY 30 LGKKGDVTELCTGS--QKSIQHFWMKSNQIKILGNG-----SFTTKPSKLNBRAD 81
DB 30 LARPGASVLSCKSGSYTLTISGISW-----VKQRTGGLEWIGETIYFGSGNSYFNEFK 84
QY 82 SRRSLW--DOGNFELIKNLKIEDSDTYICEVEDQKEVOLLVFGLTANSDTHLQOQS 138
DB 85 GKATLTVDKSSSTAYLHLSLSLSDSANVYFCA---GPRQVGLLPRG-----YNGQS 132
QY 139 LTLTLESPGSSPSVQ-----CRSPRGKNI-----QSGKTLV--- 171
DB 133 TLVTASAKTTPSPSYPLAPGCGDTTGSSTVTLGLGVKGFPESVTVWNSSGLSSVYTF 192
QY 172 -----SQLELDQSGTW-----TCTVLAQ--NQKQVEFKID-----LYPCAPE 206
DB 193 PALIQSGLYTWSSSVTVPSSTWPSQTVTCVVAHPASSTVVDKLEPSPGTSTINPC--- 249

```

```

QY 207 PKSCDKHTTC--PELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWY 264
DB 250 --PKECHCKCAPRLLEGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWY 307
QY 265 DGEVFNHAKKPREEQYNSTIRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 324
DB 308 NNVEVLTAQOTQTHREDNSTIRVSVLPIQHODMNGKEYKCKVSNKALPAPIEKTISKI 367
QY 325 KQGPPEVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPVLD 384
DB 368 KGIYRABQVYTLSPPELRSKQVSLTCLVAGSPEDISVEWISNGHTEENYKOTAPLVD 427
QY 385 SDGSFFLYSKLTVDKSRMVOGNVFGSCVMHEALHNHYTOKSLSLSPG 431
DB 428 SDGSFFLYSKLTNMKTSKMEKTDSPSCVNRHGLKNYLLTKTISRSPG 474

```

#### RESULT 38

PS0019

Ig gamma-2a chain C region - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 16-Jul-1999

C:Accession: PS0019; D25941

R:Brueggemann, M.

Gene 74, 473-482, 1988

A:Title: Evolution of the rat immunoglobulin gamma heavy-chain gene family.

A:Reference number: PS0017; MUID:89232738; PMID:3319946

A:Accession: PS0019

A:Molecule type: DNA

A:Residues: 1-322 <BR2>

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 216-322 <BR2>

C:Genetics:

A:Introns: 98/1; 109/1; 216/1

C:Superfamily: Immunoglobulin C region; immunoglobulin homology

C:Keywords: Immunoglobulin

F:20-84/Domain: immunoglobulin homology <IMM>

Query Match 29.3%; Score 792.5; DB 2; Length 322;

Best Local Similarity 51.0%; Pred. No. 4.7e-40;

Matches 160; Conservative 47; Mismatches 72; Indels 35; Gaps 8;

```

QY 133 LLOQ---QSLTLTLESPGSS-----PSVOCSPRGKNIQSGKTLVSQLELDQSGTW-- 182
DB 28 LVNGVFPPEPVTVMNNGALSSGVHTFPAL-----LQSGLYTLSSVTV--PSSTWSS 77
QY 183 ----TCTVLAQ--NQKQVEFKIDYVPCPAPEKSCDKHTCPBELLAGSPVFLFPKPKDTLM 237
DB 78 QAVTCVVAHPASSTKVAKIEPRVITQNPCC---NPGCGTGSSEV-----SSVTFPPKTKDVL 127
QY 238 ISRTPEVTCVVVDVSHEDPEVKFMWYDGVNAHAKKPREBOYNSTIRVSVLTVLHQD 297
DB 128 ITLTPKVTCCVVVDISQNDPEVRFSWFTDVEVHTAQTABEKNSTLRSELPVTHRD 187
QY 298 WLNQKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGF 357
DB 188 WLNQKTFKCKVNSGAPAPAPIEKSISKEGTFPRGQVYVTMAPKPEMNGSQVSTICWYKGF 247
QY 358 YPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRMVOGNVFGSCVMHEAL 417
DB 248 YPDIYTEMKNGQPCQENNYKTTPMTDSDGSFFLYSKLTANVKETWQGNFTFTCSVLHGL 307
QY 418 HNHYTOKSLSLSPG 431
DB 308 HNHYTOKSLSLSPG 321

```

## RESULT 39

RMWOT4

T-cell surface glycoprotein CD4 - rhesus macaque

N.Alternate names: T-cell surface antigen T4/Leu 3

C.Species: Macaca mulatta (rhesus macaque)

C.Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999

C.Accession: C32722

R.Cameron, D.; Seed, B.

Cell 60, 747-754, 1990

A.Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi

A.Reference number: A32722; MUID:90182664; PMID:2107024

A.Accession: C32722

A.Molecule type: mRNA

A.Residues: 1-432 &lt;CAM&gt;

A.Cross-references: GB:M31134

C.Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells

C.Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C.Keywords: duplication; glycoprotein; T-cell; transmembrane protein

F.1-432/Product: T-cell surface glycoprotein CD4 #status predicted &lt;MNT&gt;

F.1-371/Domain: extracellular #status predicted &lt;EXT&gt;

F.9-86/Domain: immunoglobulin homology &lt;IM1&gt;

F.111-161/Domain: immunoglobulin homology #status atypical &lt;IM2&gt;

F.180-293/Domain: immunoglobulin homology &lt;IM3&gt;

F.296-347/Domain: immunoglobulin homology &lt;IM4&gt;

F.372-395/Domain: transmembrane #status predicted &lt;TMW&gt;

F.396-432/Domain: intracellular #status predicted &lt;INT&gt;

F.16-84,130-159,303-345/Disulfide bonds: #status predicted

F.271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 29.2%; Score 790; DB 1; Length 432;

Matches 152; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

Qy 27 KYVLGKGGTVBELTCTASQKSIQFHWKNSNQIKINGOSFLTKGSPSLNDRASRL 86

Db 2 KYVLGKGGTVBELTCTASQKNTQFHWKNSNQIKINGOSFLTKGSPSLNDRASRL 61

Qy 87 WPOGNPLLIKNLKIEDSTYICEVEDKEVQLVFGILTANSDTHLGGSLTLTLESP 146

Db 62 WPOGNPLLIKNLKIEDSTYICEVEDKEVQLVFGILTANSDTHLGGSLTLTLESP 121

Qy 147 PSSSPVQCRSPRGKNIQGGKTLVSQLELDGSGTCTVTLNOKKVEFKIDIV 200

Db 122 PSSSPVQCRSPRGKNIQGGKTLVSQLELDGSGTCTVTLNOKKVEFKIDIV 175

## RESULT 40

S06611

Ig gamma-2 chain C region (clone 32.2) - bovine (fragment)

C.Species: Bos primigenius taurus (cattle)

C.Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 16-Jul-1999

C.Accession: S06611; B31303

R.Symons, D.B.A.; Clarkson, C.A.; Beale, D.

Mol. Immunol. 26, 841-850, 1989

A.Title: Structure of bovine immunoglobulin constant region heavy chain gamma 1 and gamma

A.Reference number: S06610; MUID:90087956; PMID:2513487

A.Accession: S06611

A.Molecule type: DNA

A.Residues: 1-327 &lt;SYM&gt;

A.Cross-references: EMBL:X16702

A.Note: the sequence was determined from the germline gene

C.Genetics:

A.Gene: Ig CH gamma 2

A.Introns: 99/1; 112/1; 219/1

C.Superfamily: immunoglobulin C region; immunoglobulin homology

C.Keywords: glycoprotein; immunoglobulin; membrane protein

F.20-85/Domain: immunoglobulin homology &lt;IMW&gt;

F.175/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 28.6%; Score 774; DB 2; Length 327;

Matches 146; Conservative 41; Mismatches 55; Indels 14; Gaps 4;

Qy 181 TWTCTVLO--NOKVEFKIDI-VPCPAPEPKSCDTHTCPELLGSPVFLFPPEKIDITLM 237

Db 80 TETCVAPPASTKTDKAVGSSDCSKRNQNCVRE-----PSVFIFFPKKIDITLM 130

Qy 228 ISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVNAKTKPREQYNSTYRVSVLTVLIQD 297

Db 121 ITGTEPVCVVVNVGHNDPEVQFSWFVDVDEVHTARTKPREQYNSTYRVSVLTQIHOD 190

Qy 298 WINGEYCKVSNKMLPAPIETISKAGTPEPPQVTLTPSRDELTKNQVSLTCLVKGF 357

Db 191 WNGGKEFCKVKNIKGLSASVIAIISRSKGPAPPEPVYLDPEKEISKVSLTQVIGF 250

Qy 358 YPSDIAVEMESNGQ--PENNYKTTPVLDSDGSPFLYSKLTVDKSRMOQGNFSCSVME 415

Db 251 YPEDVDVEMKQRRORESEDKYTTTPOLDADNSYFLYSKLRVDRNSWQRGDYTCVME 310

Qy 416 ALHNHYTKSLSLSPG 431

Db 311 ALHNHYTKSLSKSG 326

## RESULT 41

I46732

Ig gamma heavy chain constant region - rabbit (fragment)

C.Species: Oryctolagus cuniculus (domestic rabbit)

C.Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 21-Jan-2000

C.Accession: I46732

R.Heldmann, O.; Rougeon, F.

Nucleic Acids Res. 10, 1535-1545, 1982

A.Title: Molecular cloning of rabbit gamma heavy chain mRNA.

A.Reference number: I46732; MUID:82174328; PMID:6280149

A.Accession: I46732

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-180 &lt;HEI&gt;

A.Cross-references: GB:J00665; NID:9165109; PIDN:AAA31288.1; PID:9165110

C.Superfamily: immunoglobulin C region; immunoglobulin homology

F.93-160/Domain: immunoglobulin homology &lt;IMW&gt;

Query Match

Best Local Similarity 26.0%; Score 702; DB 2; Length 180;

Matches 124; Conservative 26; Mismatches 28; Indels 0; Gaps 0;

Qy 254 EDPPEKFMVYDVGVEVNAKTKPREQYNSTYRVSVLTVLIQDNLNGEYCKVSNKML 313

Db 2 DDPVQFTWYINNEQVARTAPPLRQDQFNSTIRVSTLPIAQDWLRGEFCKVHNKAL 61

Qy 314 PAPIETISKAGTPEPPQVTLTPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPE 373

Db 62 PAPIETISKAGTPEPPQVTLTPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPE 121

Qy 374 NNKYKTPPVLDSDGSPFLYSKLTVDKSRMOQGNFSCSVMEALHNHYTKSLSLSPG 431

Db 122 DNYKTTTPVLDSDGSPFLYSKLTVDKSRMOQGNFSCSVMEALHNHYTKSLSLSPG 179

## RESULT 42

A46254

CD4 precursor - rabbit

C.Species: Oryctolagus cuniculus (domestic rabbit)

C.Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000

C.Accession: A46254

R.Hague, B.F.; Sawasdi-kosol, S.; Brown, T.J.; Lee, K.; Recker, D.P.; Kindt, T.J.

Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992

A.Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency vi

A.Reference number: A46254; MUID:92290370; PMID:1518821

A.Accession: A46254

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-459 &lt;HAG&gt;

A.Cross-references: GB:M92840; NID:9164871; PIDN:AAA31198.1; PID:9164872

A.Note: sequence extracted from NCBI backbone (NCBIN:112732, NCBI:112733)

C.Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

F:322-372/Domain: immunoglobulin homology <IMM>

Query Match 22.3%; Score 601.5; DB 2; Length 459;  
 Best Local Similarity 41.9%; Pred. No. 1,66-28;  
 Matches 153; Conservative 47; Mismatches 98; Indels 67; Gaps 11;

OY 1 MNRGVFPHLLLVQLALLPAATQGNKVLGKKGDTVELTCTAQQKSIOFHMKNSNQIK 60  
 DB 1 MNRRIYFQCLLVLEFLLPAPATMGKTAVRGKAGAIIVELPCOSSQKRNSVENMKIANQVK 60  
 OY 61 ILNGG-----SFLTGPBKNDRAISRSLMDQGFPLIKLKIEDSDTYICEVEDQKE 116  
 DB 61 ILNGGSSSSSSFWLKGNSPLNSRVESKKNMDDQGFPLVIKDLRRDDGTICVCEGDCKM 120  
 OY 117 EVOLLVFLGLTANSDTHLLQGSLLTLLESPPSSPVOCSPRGKNIQGGKTLVSQLEL 176  
 DB 121 EVELLVFLTLNPNTRLHGOSLTLTLGSPVGSVVMKSPENKIITGFTCSMPKRL 180  
 OY 177 QDSGTWCTV-LQNGKVEFKIDIVPCAPBPCKSCDKTHTCPELIGSPVFLFPKPKDT 235  
 DB 181 QDSGTWSCHLSFQDNKLELDIKIIVLGFPKASA-----TYKKKEGEVERSF----- 229  
 OY 236 LMISTPEVTGVVDVSHEDPEV--KFWWYVDGVEVHNAKTKPREEQYNSTRVVSULTV 293  
 DB 230 -----LNFEDESLSGELMWQVDGAS----- 249  
 OY 294 LHQDWLNGKYEKKVKS-NKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLT 351  
 DB 250 SAQSVSVSLSDRKVSQKILP--DLKIQMSKILPLS---LTLFQALHRYAGSNLSLT 303  
 OY 352 CLVNG 356  
 DB 304 -LDKG 307

#### RESULT 43

S30193  
 T-cell surface glycoprotein CD4 - dog  
 C/Species: Canis lupus familiaris (dog)  
 C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 21-Jul-2000  
 C/Accession: S30193  
 R/Milte, K.F.; Conner, G.E.; Mitcz, D.H.; Alejandro, R.  
 Biochim. Biophys. Acta 1172, 315-318, 1993  
 A/Title: Primary structure of the canine CD4 antigen.  
 A/Reference number: S30193; MUID:93192324; PMID:791632  
 A/Accession: S30193  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-432 <ML>  
 A/Cross-references: EMBL:X68565; NID:q288652; PIDN:CA837664.1; PID:94467377  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: glycoprotein  
 F:202-311/Domain: immunoglobulin homology <IMM>

Query Match 21.5%; Score 581.5; DB 2; Length 432;  
 Best Local Similarity 32.8%; Pred. No. 2,36-27;  
 Matches 170; Conservative 67; Mismatches 152; Indels 129; Gaps 19;

OY 12 LVQLALLPAATQGNKVLGKKGDTVELTCTAQQKSIOFHMKNSNQIKILNGGSPFTK 71  
 DB 1 LMLDLVMLPAVTPVAREVVLGKAGDAVELPCQTSQKNIHFMRDSSMWQIILNGGSEFTV 60  
 OY 72 GPKLNDRAISRSLMDQGFPLIKLKIEDSDTYICEVEDQKE 127  
 DB 61 GSSSLKRVESKKNIMDDQGFPLVIKDLRVDSDGTYFCDT-DKQEVLLVFNLTAKKDS 119  
 OY 128 -----NSDTHLLQGSLLTLLESPPSSPVOCSPRGKNIQGGKTLVSQLELQDSGTW 182  
 DB 120 GSSGSSNIRLLQGGQLTLTLLENPSGSSPVQMGPKMSHGQGNLSLSPLELQDSGTW 179  
 OY 183 TCTVLQNGKVEFKIDIVPCAPBPCKSCDKTHTCPELIGSPVFLFPKPKDTLMISRT 242  
 DB 180 TCTISQKVEFMINVLVLAFA--QKVSNTFYARE--GDQVEFSF----- 221

OY 243 EVTCVVVDVSHEDPEV--KFWWYVDGVEVHNAKTKPREEQYNSTRVVSULTV LHQDWLN 300

DB 222 -----LSEEDNVLGELRMQNGAS-----SSLL-----WIS 248

OY 301 GKEYCKVSNKALPAPIEKTISKAKGPRE--POVYTLPPSRDELTKNOVSLTCLVNGFY 358

DB 249 FTLENRKLSMKEAHAPL-----KLQMKESLPRLFPQLSRVAGSGIITLNLAKGTL 301

OY 359 PSDIAVEMESNGQENNYKTPPLVDSGSEFLYSKLTVDKSRMQGVFSCVMEKYLH 418

DB 302 YQEV-----NLVMEANSSQNNL-TCEVLGP--- 326

OY 419 NHYQKSLSPGLQDETCAEADGELDGLMTDPPRASALPAPFGSALPDPQ---TA 475

DB 327 ---TSPLETLIS--LNLEQAAKVK--QOKLVWVVDPEEGT-----NQCILSDXDKVILA 374

OY 476 SAL--PDPASALPALA-----VISFLGLGLGVAC 506

DB 375 SSLVSSPVVIKSWPKFLATTLGILGLLLIGLCVFC 412

#### RESULT 44

A36040  
 Ig heavy chain V-III region (ART) - human (fragments)  
 C/Species: Homo sapiens (man)  
 C/Date: 16-Nov-1990 #sequence\_revision 16-Nov-1990 #text\_change 21-Jan-2000  
 C/Accession: A36040  
 R/Bultz, M.; Weiss, D.T.; Solomon, A.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 6542-6546, 1990  
 A/Title: Immunoglobulin heavy-chain-associated amyloidosis.  
 A/Reference number: A36040; MUID:90370821; PMID:2118650  
 A/Accession: A36040  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 1-218 <EUL>  
 C/Superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: heterotetramer; immunoglobulin  
 F:132-199/Domain: immunoglobulin homology <IMM>

Query Match 21.3%; Score 574.5; DB 2; Length 218;  
 Best Local Similarity 77.5%; Pred. No. 2,66-27;  
 Matches 110; Conservative 3; Mismatches 12; Indels 17; Gaps 1;

OY 290 VLTGVLDWLNNGKYEKKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVS 349  
 DB 94 VAAALLTDWNGG-----FLVTYSSQPREPOVYTLPPSRDELTKNOVS 136  
 OY 350 LTCLVKGFPSPDIIVEMESNGQENNYKTPPLVDSGSEFLYSKLTVDKSRMQGVNFS 409  
 DB 137 LTCLVKGFPSPDIIVEMESNGQENNYKTPPLVDSGSEFLYSKLTVDKSRMQGVNFS 196  
 OY 410 CSVMEHLNHYTQKSLSPG 431  
 DB 197 CSVMEHLNHYTQKSLSPG 218

#### RESULT 45

S69340  
 Ig heavy chain VHIII-D-JH-CH3 region - human  
 C/Species: Homo sapiens (man)  
 C/Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 21-Jan-2000  
 C/Accession: S69340  
 R/Khamlich, A.A.; Ancourtier, P.; Preud'homme, J.L.; Cogne, M.  
 Eur. J. Biochem. 229, 54-60, 1995  
 A/Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
 A/Reference number: S69339; MUID:95262687; PMID:7744049  
 A/Accession: S69340  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-249 <KHA>  
 A/Cross-references: EMBL:X81696  
 C/Superfamily: immunoglobulin C region; immunoglobulin homology

F1162-229/Domain: immunoglobulin homology <IMM>

Query Match 21.2%; Score 572.5; DB 2; Length 249;

Best Local Similarity 62.9%; Pred. No. 4e-27; Indels 37; Gaps 3;

Matches 117; Conservative 9; Mismatches 23; Indels 37; Gaps 3;

266 GVEVNAKTPREEDQNSTYRVSVLT-----VLHDMWNGKEYCKVSN 310

80 GDSVGRFTISRDNSNOLYLOMNSTLADTAVYCAKGVSVIH----- 124

311 KALPAPIE-----KTSKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 365

125 --LPSAIFHWGQSTRYIVSSGQPREPOVYTLPPSRREEMTKNQVSLTCLVKGFYPSDIAVE 182

366 WESNQPENNYKTPPVLDSDGSFLYSKLTLYDCKSNQGNVFCGSVMHEALHNYTQKS 425

183 WESNQPENNYKTPPVLDSDGSFLYSKLTLYDCKSNQGNVFCGSVMHEALHNYTQKS 242

426 LSLSPG 431

243 LSLSPG 248

RESULT 46

S14236

IG gamma-1 chain C region (15C5) - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 25-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 23-Jul-1999

C/Accession: S14236

R/Vandamme, A.M.; Bulten, F.; Bernar, H.; Nelles, L.; Lijnen, R.H.; Collen, D.

Eur. J. Biochem. 192, 767-775, 1990

A/Title: Construction and characterization of a recombinant murine monoclonal antibody

A/Reference number: S14236; MUID:91006173; PMID:2209622

A/Accession: S14236

A/Molecule type: mRNA

A/Residues: 1-152 <VAV>

A/Cross-references: EMBL:X56393; NID:951617; PIDN:CAA39804.1; PID:951618

C/Superfamily: immunoglobulin C region; immunoglobulin homology

C/Keywords: immunoglobulin

F165-132/Domain: immunoglobulin homology <IMM>

Query Match 20.9%; Score 566; DB 2; Length 152;

Best Local Similarity 61.6%; Pred. No. 5.4e-27;

Matches 93; Conservative 33; Mismatches 25; Indels 0; Gaps 0;

281 YNSTRVSVSLTVLHDMWNGKEYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSR 340

1 FNSTRVSVSLTVLHDMWNGKEYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSR 60

341 DELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFLYSKLTLYDCK 400

61 EQMADKSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFLYSKLTLYDCK 120

401 RWQGNVFCGSVMHEALHNYTQKSLSLSPG 431

121 NWEAGTFTCSVLHDEGLHNHHTEKSLSHSPG 151

RESULT 47

A27449

T-cell surface glycoprotein CD4 precursor - rat

N/Alternate names: M3/25 antigen

C/Species: Rattus norvegicus (Norway rat)

C/Date: 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 21-Jan-2000

C/Accession: A27449; A35433

R/Clark, S.J.; Jeffries, W.A.; Barclay, A.N.; Gagnon, J.; Williams, A.F.

Proc. Natl. Acad. Sci. U.S.A. 84, 1649-1653, 1987

A/Title: Peptide and nucleotide sequences of rat CD4 (M3/25) antigen: evidence for deriv

A/Reference number: A27449; MUID:87175535; PMID:3104900

A/Accession: A27449

A/Molecule type: mRNA

A/Residues: 1-457 <CLAV>

A/Cross-references: GB:M15768; NID:9203387; PIDN:AAA0901.1; PID:9203388

R/Davis, S.J.; Ward, H.A.; Puklavac, M.J.; Willie, A.C.; Williams, A.F.; Barclay, A.N.

J. Biol. Chem. 265, 10410-10418, 1990

A/Title: High level expression in Chinese hamster ovary cells of soluble forms of CD4 T

A/Reference number: A35433; MUID:90285164; PMID:2113054

A/Contents: annotation

C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C/Keywords: glycoprotein; membrane protein; surface antigen

F219-300/Domain: immunoglobulin homology <IMM>

Query Match 18.3%; Score 495; DB 2; Length 457;

Best Local Similarity 32.8%; Pred. No. 3.3e-22;

Matches 137; Conservative 57; Mismatches 132; Indels 92; Gaps 14;

1 MNRGPPRHL--LVLQIALLPAAQGNKVVLTGKGVDELCTASQKKSIOFHKNNSQ 58

1 MCRGSPFHLPLRLLQSLKLVVQKTVLVGKGSALPCBSTSRSSAFWAKSDQ 60

59 IKILGNQSFLLTKGSPKNDRASSRSLMDQGNFLIKNLKIEDSDTYICEVEDQKEV 118

61 KTLIGYKMKLLIKGSLFLYSRDSKKNAMERGSFLLINKLRMEDSQTLYCELENKKEV 120

119 QLVFGLTNSDTHLLQGSLLTLES--PSSSPSVQCRSPRGKNIQGGKTLVSQLEIQ 177

121 ELWVFRVTFNPQTRLLQGSLLTLLDSNPKVSDPIECKKSSNIVKDSKAFSTSLRIQ 180

178 DSGTCTVVLQNKQKVER--KIDIVPCAPAPKSCDKHTHCELLGFSVLPFPKPKDT 235

181 DGINCTVTLNOKHSDPMKLSVL-----GFASTITAYKSGESAESFP----- 227

236 IMISRPVTCVVDVSHEDPEKFNWYVDGEVNAKTKPREEDQNSTYRVSVSLTVLH 295

228 -----LNLGESSLQELRW-----KAKKAPSS----- 249

296 QDMWNGKEYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNQVSLTCLVK 355

250 QSMITFSLKQKVS-----VQKSTSNPKFQLSR-----TLQI- 285

356 GFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFLYSKLTLYDCKSNQGNVFCGSVM 413

286 ----QVSLQFAGSG---NLTLT---LDR-GILVEVNLVVMKVQDPDSNLTCEVM 331

RESULT 48

RMKST4

T-cell surface glycoprotein CD4 precursor - mouse

N/Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3

C/Species: Mus musculus (house mouse)

C/Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999

C/Accession: A02110; A26038; A39893; A39955; I54664; I69018; A47642

R/Gorman, S.D.; Gorman, S.D.; Field, E.H.; HunnKapiller, T.; Parne, J.R.

Science 234, 610-614, 1986

A/Title: Isolation and sequence of L3T4 complementary DNA clones: expression in T cells

A/Reference number: A02110; MUID:87018845; PMID:3094146

A/Accession: A02110

A/Molecule type: mRNA

A/Residues: 1-457 <TOU>

A/Cross-references: GB:M13816; NID:9192070; PIDN:AAA37267.1; PID:9309112

R/Littman, D.R.; Gettner, S.N.

Nature 325, 453-455, 1987

A/Title: Unusual insertion in the immunoglobulin domain of the newly isolated murine CD4 (L

Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987

A/Reference number: A39893; MUID:88041159; PMID:2823269

A/Accession: A39893

A/Molecule type: DNA

A/Residues: 1-25, 'E' 27-457 <GOR>

A/Cross-references: GB:M17080; GB:J03003; NID:9192515; PIDN:AAA37402.1; PID:9387124



R.Maddon, P.J.; Molineaux, S.M.; Maddon, D.E.; Zimmerman, K.A.; Godfrey, M.; Alt, F.W.;  
 Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987  
 A:Title: Structure and expression of the human and mouse T4 genes.  
 A:Reference number: A39955; MUID:88097446; PMID:3501122  
 A:Accession: A39955  
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 25-457 <MAD>  
 A:Note: the cited Genbank accession number, J03564, is not in release 101.0  
 R.Parnes, J.R.; Hunkapiller, T.  
 Immunol. Rev. 100, 109-127, 1987  
 A:Title: L3T4 and the immunoglobulin gene superfamily: New relationships between the im  
 A:Reference number: 154564; MUID:88152875; PMID:3326818  
 A:Accession: 154564  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-457 <RES>  
 A:Cross-references: GB:M36850; NID:G198670; PIDN:AAA39401.1; PID:G198671  
 A:Accession: 159018  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 208-318 <RE2>  
 A:Cross-references: GB:M36851; NID:G198672; PIDN:AAA39402.1; PID:G554183  
 R.Clackson, B.U.; Isagatoros, J.; Kitzbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.; M  
 Immunogenetics 23, 129-132, 1986  
 A:Title: The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.  
 A:Reference number: A47642; MUID:8616694; PMID:3082751  
 A:Accession: A47642  
 A:Molecule type: protein  
 A:Residues: 27-43 <CLA>  
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells  
 C:Genetics:  
 A:Introns: 18/1; 74/1; 128/1; 207/1; 319/1; 386/1; 425/3; 448/2  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: alternative initiators; duplication; glycoprotein; T-cell; transmembrane pro  
 F:1-26/Domain: signal sequence #status predicted <SIG>  
 F:27-457/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>  
 F:35-114/Domain: immunoglobulin homology <IM1>  
 F:139-190/Domain: immunoglobulin homology <IM2>  
 F:220-301/Domain: immunoglobulin homology <IM3>  
 F:241-457/Product: CD4, brain-specific short form #status predicted <BRA>  
 F:321-372/Domain: immunoglobulin homology <IM4>  
 F:395-419/Domain: transmembrane #status predicted <TM>  
 F:420-457/Domain: intracellular #status predicted <INT>  
 F:42-112, 159-188, 328-370/Diulfide bonds: #status predicted  
 F:187, 298, 323, 392/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.6%; Score 475; DB 1; Length 457;  
 Best Local Similarity 52.8%; Pred. No. 5, 1e-21;  
 Matches 104; Conservative 32; Mismatches 55; Indels 6; Gaps 5;

Oy 1 MNRGVPRFH-LLVLTALLPAAQGNVVGKKGDVTELTCTASOKKSIQFHKNSQI 59  
 Db 1 MCRALSRRLLLLLQLAVTQKTLVKGKGBSELCESSOKKTIYFTWKFSQR 60

Oy 60 KILNQG-SPLTKG-PSKLNDRADSRSLDQGNFPLIKNLKIEDSTYICEVEDQKE 116  
 Db 61 KILQHGKGVILIRGSSPSQF-DRFDKSKGAKWKSFPILIKLMEBSQTYICLENRKE 119

Oy 117 EVOLLVRELTLNSTDHLQGSLLTLES-PPGSSPSVQCSPPGKNIQSKTLSSVQLE 175  
 Db 120 EVELLVFVETPSPTSLQGSLLTLDLSNKSVPNPLTECHGKXKVSQSKVLSMNL 179

Oy 176 LODSGTWCTVTLONOKK 192  
 Db 180 VQSDSPMNCVTTLDOCK 196

RESULT 49  
 B46529  
 Ig Y heavy chain (7.8S) - duck  
 N:Alternate names: Ig gamma chain (7.8S)  
 C:Species: Anas platyrhynchos (domestic duck)

C:Date: 18-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
 A:Accession: B46529; S20759  
 R:Magor, K.E.; Marr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.  
 J. Immunol. 149, 2627-2633, 1992  
 A:Title: Structural relationship between the two IY of the duck, Anas platyrhynchos: mo  
 A:Reference number: A46529; MUID:93017865; PMID:1401901  
 A:Accession: B46529  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-572 <MAG>  
 A:Cross-references: EMBL:X65219; NID:962442; PIDN:CAA6322.1; PID:962443  
 A:Experimental source: spleen  
 A:Note: sequence extracted from NCBI backbone (NCBI:P.116127)  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:37-120/Domain: immunoglobulin homology <IM>

Query Match 15.2%; Score 411.5; DB 2; Length 572;  
 Best Local Similarity 30.0%; Pred. No. 4e-17;  
 Matches 103; Conservative 55; Mismatches 136; Indels 49; Gaps 10;

Oy 102 EDSPTVCEV---EDQKEVQLVFGLTANSDDLQGSLLTLESPPGSSPSVQCRSP 158  
 Db 265 DDSVELLCVITGSPSPPEVEWLDGAPAH-----LVATMRPQ-----REA 306

Oy 159 RGNKIQGKTLSSVQLELQDSGTWCTVQLQ-----NQKVEFKIDIVCPAPAPKSCDK 212  
 Db 307 GSKTYMATSTNVNSREWKAGKAFCTCKVNPATGCTGQGNARF-----CRGSAQSCSP 360

Oy 213 THTCPPELLGSPSVFLPFPKPKDTLMTSRPEVTCVVVDVSHEDPEVKFNMYDGEVANA 272  
 Db 361 I-----QIFVVPSP-GSLYIRQDAKIVCLVNLPSDASLSISWTRKSGALRP 408

Oy 273 KTKPREQVSTRVSVLVTLHODMLNGEKYCKNSKALPPIETKISKAKQPREPQ 332  
 Db 409 DPMVLTEHFGTFTASSSLAISTQDWLAGRFCTVQHEDLPPVLGKSIAGAKVTAPY 468

Oy 333 VYTLPSRDLTLTKQVSLTCLVKGFPYPSDIAVEW--ESNQPENNYKTPPVLD--SDGS 388  
 Db 469 IFFPFPRAEELSLAEVLTCLVKGFPQENHVEVQMLRNHNVPAAEFYTTTPPLKPNDDGT 528

Oy 389 FFLYSKITLVDSKRWQGNVFSVMEHAIHNYTQSLSPG 431  
 Db 529 FFLYSKITVPKASWQGVSYACVMEHGLPMRFQRLQKTPG 571

RESULT 50  
 S04845  
 Ig heavy chain precursor - African clawed frog (Xenopus laevis)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 26-Aug-1999  
 A:Accession: S04845; S05695  
 R:Amemiya, C.T.; Haire, R.N.; Litman, G.W.  
 Nucleic Acids Res. 17, 5388, 1989  
 A:Title: Nucleotide sequence of a cDNA encoding a third distinct Xenopus immunoglobulin  
 A:Reference number: S04845; MUID:89345103; PMID:2503814  
 A:Accession: S04845  
 A:Molecule type: mRNA  
 A:Residues: 1-549 <AME>  
 A:Cross-references: EMBL:X15114  
 R:Litman, G.W.  
 submitted to the EMBL Data Library, April 1989  
 A:Reference number: S05695  
 A:Accession: S05695  
 A:Molecule type: mRNA  
 A:Residues: 'LC', 3-308, 'H', 310-549 <LIT>  
 A:Cross-references: EMBL:X15114; NID:964799; PID:G763031  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: glycoprotein; heterotrimer; immunoglobulin  
 F:26-109/Domain: immunoglobulin homology <IM>  
 F:281, 294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 14.9%; Score 402.5; DB 2; Length 549;

Best Local Similarity 29.1%; Pred. No. 1.3e-16;  
Matches 118; Conservative 65; Mismatches 134; Indels 89; Gaps 17;

QY 67 WQGG-----NPLIK-----NKIEDSD-----TYICEVE----- 112  
DB 171 WNGSITSGLNKPPALQSGSLPASSQLTILPSDWKAKKSPCEVNEHKPTSTKYQKIE 230  
QY 113 --DOKEVO--LTVFGLTANSDT-----HLLQGSLTLLTLESPGS 149  
DB 221 CDDEPEPIEPTVEIILQGPASSKSVELLLITGVASEIKYMWLNGQ---VTNISPSNS 287  
QY 150 SPVSQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLQNOQKVEFKIDIVCPAPEPKS 209  
DB 288 KP---CKEENG--TFSSRSKVSVPKEDMNSDSYTCVTH-----PASHTKT 329  
QY 210 CDKHTTCPELLGGPSVFLFPPPKDITLMSRTEPVTGVVVDVSHDEPVKFW--YVDGE 268  
DB 330 EASTKKCDTATTPKVDVLPSPKD--LVYTKAKYCVISMASTD--DLTVQMSRSDGK 387  
QY 269 VNAKTKPREQYNSTYRVSVLTVLHODMLNGKEKCVSNKALPAPIETKISAKGP 328  
DB 388 ALAPDSAP--EKAYDGTFTYKSLTKISPGDMENKQPCNVVHPDLPSLEKSIQKSDPG 446  
QY 339 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNG--QPENNYKTTPPYLDSD 386  
DB 447 TEPITLPPSDDELRLNDFISLICMLKNRPQDIVFWMKQGVTLIEDYVMTTTPVLEBE 506  
QY 387 GSFFL-YSLUTYDKSRMQGNVFGSCVWHEALHNHYTQKSLSPGLQDETCAEADGEL 431  
DB 507 EBGFTSFSKLTITARSQWNRGATYSCI---AAHNTISORDIKNRG 548

## RESULT 51

S14683  
Ig mu chain precursor, membrane-bound (clone 201) - human  
C/Species: Homo sapiens (man)  
C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 23-Jul-1999  
C/Accession: S14683; S08047  
R.Friedlander, R.M.; Nusseizweig, M.C.; Leder, P.  
Nucleic Acids Res. 18, 4278, 1990  
Article: Complete nucleotide sequence of the membrane form of the human IgM heavy chain.  
A/Reference number: S14683; MUID:90332450; PMID:2115996  
A/Accession: S14683  
A/Molecule type: mRNA  
A/Residues: 1-627 <FRI>  
A/Cross-References: EMBL:X17115; NID:g33450; PIDD:CAA34971.1; PID:g33451  
C/Superfamily: immunoglobulin C region; immunoglobulin homology  
C/Keywords: immunoglobulin; membrane protein  
F.1-15/Domain: signal sequence #status predicted <Sig>  
F.16-627/Product: Ig mu chain #status predicted <MAT>  
F.34-117/Domain: immunoglobulin homology <IMM>

Query Match 14.4%; Score 389; DB 2; Length 627;  
Best Local Similarity 23.9%; Pred. No. 9.6e-16;  
Matches 116; Conservative 77; Mismatches 167; Indels 126; Gaps 16;

QY 36 TWELTCTASQ--KKSQIFWKNKSNQIKILGNQSFITKGPSKLNDRADSRSLMQGNFP 93  
DB 175 SVAVCCGLADPLPDSITFSMKYKNNSDISSTRG-----FP 209  
QY 94 LIIKULKIEDSDTYICEVEDOKEEVQLVFGGLTANSPTLL-----QGSLTLLT 142  
DB 210 SVLRGKKAATSQVLLPSKD-----VMQGTDEHVCKVQHPNGKNKKNVPLPYI 258  
QY 143 LESPSSPSVQCR-----SPR-----GKNIQGGKTLVSQ 173  
DB 259 AELPFRVSVFVPRDQFPGNPRSKSLICQATGFSRQIQVSWLRBGQVGSVTTDQVQ 318  
QY 174 LELQDSG-----TWICTVLQNOQKVEFKIDIVCPAPEPKS 212  
DB 319 ABAKESGPTTYKVTSTLTIKESDWLSQSMFTCRV--DHKGLTFQGNASSMCVPPDDTAIR 376  
QY 213 THTCELLGGPSVFLFPPPKDITLMSRTEPVTGVVVDVSHDEPVKFWVYVDGEVHNA 272

DB 377 -----VFALPPS--FASIFLTSTLTCLVLTLYTD--SVTISWTRONGEAVKT 422  
QY 273 KTKPREQYNSTYRVSVLTVLHODMLNGKEKCVSNKALPAPIETKISAKGP--REP 331  
DB 423 HTNISSEHPNMTFSAVNGASICEDWNSGERFTCVTHTDLPSPKQTIISRKGVALLRP 482  
QY 332 QVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQ--ENNYKTTPPYLD--SD 386  
DB 483 DVIILPPARQNLRESATITCLVYGFSPADYFQVMQGRQPLSEKVTASMPPEPQAP 542  
QY 387 GSFFLYSLUTYDKSRMQGNVFGSCVWHEALHNHYTQKSLSPGLQDETCAEADGEL 446  
DB 543 GRFYASLITVSEEBWMTGETYTCVVAHEALPNRYTERTVKSS-----TEGVSADERGF 597  
QY 447 DGLWTT 452  
DB 598 ENLMAT 603

## RESULT 52

EMMS  
Ig epsilon chain C region (version 1) - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 16-Jul-1999  
C/Accession: A02144  
R.Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.  
Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982  
Article: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.  
A/Reference number: A02144; MUID:83117774; PMID:6818553  
A/Accession: A02144  
A/Molecule type: mRNA  
A/Residues: 1-388 <LIU>  
A/Cross-References: GB:J00476; NID:g194875; PIDD:AAA38085.1; PID:g387220  
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la C/Superfamily: immunoglobulin C region; immunoglobulin homology  
C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F.1-44/Domain: immunoglobulin homology (fragment) <IM1>  
F.81-149/Domain: immunoglobulin homology <IM2>  
F.186-254/Domain: immunoglobulin homology <IM3>  
F.290-361/Domain: immunoglobulin homology <IM4>  
F.10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 14.4%; Score 388; DB 1; Length 388;  
Best Local Similarity 28.3%; Pred. No. 6.2e-16;  
Matches 119; Conservative 77; Mismatches 138; Indels 86; Gaps 20;

QY 38 ELTCTASQKKSQIFWKNKSNQIKILGNQSFITKGPSKLNDRADSRSLMQGNPLIK 97  
DB 22 ELKVTTSQVTS---WGSKAK---NFTCHVTHPPSFNESHRT-----ILVR 59  
QY 98 NUKIEDSDTYI---CEVEDOKEEVQL--LVFGLTAN--SPTHLQGSLLTLLTLESPGS 149  
DB 60 PVNIEPTLELHSSCDPNAFHSTIQLYCFIYGHILNDVSVSWLMDNREITDTL----- 113  
QY 150 SPVSQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLQNOQKVEFKIDIVCPAPEPKS 209  
DB 114 AQTVLIKE--EGKLASTGSKNLITEQOMMSSESTFTCKV--TSQYDVLYLHTRRCPDHBR- 169  
QY 210 CDKHTTCPELLGGPSVFLFPPPKDITLMSRTEPVTGVVVDVSHDEPVKFWN----- 261  
DB 170 -----GVITVLIIPSPDL--LYQNGAPKLTCLVVDSESKKNVNVNMQEKTSV 216  
QY 262 ---WYVDGEVHNAKTKPREQYNSTYRVSVLTVLHODMLNGKEKCVSNKALPAPI 317  
DB 217 SASQVY---TKHN-----NATISLISILPVAWKWIEGYGVCYIDHDPFRKPI 263  
QY 318 EKTISKANGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQENN-- 375  
DB 264 VRSITKTPGORSAPAEVYVFPPEER--SEDKRTLTCLIQNFPEPDISVQMLDQKLISSQ 322  
QY 376 YKTPPPVLDSSG---FLYSLUTYDKSRMQGNVFGSCVWHEALHN--HYQKSLSLSPG 431

Db 323 HSTTTP-LKNSGNSGFFIFSRLEVAKTLMTQKQFTCOVHIALQKPKLEKTISSIG 381

## RESULT 53

S15590  
Ig heavy chain - human  
C/Species: Homo sapiens (man)  
C/Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C/Accession: S15590  
R/Neale, G.A.M.; Kitchingman, G.R.  
Nucleic Acids Res. 19, 2427-2433, 1991  
A/Title: mRNA transcripts initiating within the human immunoglobulin mu heavy chain enhancer  
A/Reference number: S15590; MUID:91252286; PMID:1904154  
A/Accession: S15590  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-474 <NEA>  
A/Cross-references: EMBL:X58529  
A/Note: the authors translated the codon CAA for residue 265 as Glu  
C/Superfamily: immunoglobulin C region; immunoglobulin homology  
C/Keywords: heterotrimer; immunoglobulin  
F/344-415/Domain: immunoglobulin homology <IMM>

Query Match 14.3%; Score 386.5; DB 2; Length 474;  
Best Local Similarity 23.7%; Pred. No. 9.7e-16;  
Matches 115; Conservative 78; Mismatches 167; Indels 125; Gaps 16;

QY 36 TVELTCTASQ--KKSIOFHKNKSNQIKILGNQGSFLTGPBKNDRADSRSLMDQGNFP 93  
DB 23 SVAAGCLADPFLPDSITFSWKYKNNSDISSTRG-----FP 57  
QY 94 LIINKLIKIEDSDTYICEVEDQEEVQLVFGLTANSDTLHLOGOSLTLTLESPPGS 142  
DB 58 SVLRGKRYAATSOVLTPSKDVMQGTDEHVCVKQHPGNKKNVPLPVIAELPPKV 106  
QY 143 LESPPGSSPVQCR-----SPR-----GKNIOGKTLSTLSQLELQDSC- 174  
DB 107 AELPPKAVFVPRDGFNGPRKSKLICQATGSPRQIOVSWLRGKQVSGVTTDQVQA 166  
QY 175 ELQDSC-----TWCTVLOKOKVEPKIDVPCAPPEKSCDKT 213  
DB 167 EAKSGPTTYVSTLTIKESDMLSQSMFTCRV--DHKGLTFQONASSMCPVDDTAIR- 223  
QY 214 HTCELLGSPVFLPFPKPDLMISRTPEVTCVVDVSHEDPEVKENWYDVGEVNAK 273  
DB 224 -----VFALPPS-FASIFLTSTKTLCLVTLDTLTYD-SVTISMTQNGQAVKTH 270  
QY 274 TKPREEOYNTYRVVSVLTJLHODMLNGKEYKCKVSNKALPAPIEKTISSAKGQP-REPQ 332  
DB 271 TNISSEHNATFSAVGEASICEEDMNSGERFTCTVHTDLPSPKQITISRPKGVALLRPD 330  
QY 333 VYTLPPSDELT-KNOVSLTCLVKGFPSPDIIVAVESNGQP--ENNYKTTTPVLD--SDG 387  
DB 331 VYLLPPAEQNLRESATITCLVTGFPADVFVOMQORGPLSEKVTYSAFMEPPAPG 390  
QY 388 SFPLYSKLTIVKSRMOQGNVSCSVMEALHNHYTOKSLSPGLQDTECAEQDDEL 447  
DB 391 RYFAHSITLTVSEEMNGETTYTCVAHAELPNRTERTVDSK-----TEGVSADSESGFE 445  
QY 448 GLWTT 452  
DB 446 NLMAT 450

## RESULT 54

S37768  
Ig mu chain C region - human  
C/Species: Homo sapiens (man)  
C/Date: 12-Feb-1998 #sequence\_revision 12-Feb-1998 #text\_change 23-Jul-1999  
C/Accession: S37768  
R/Hatlinratnach, N.; Donadel, G.; Sigounas, G.; Nockins, A.L.  
Mol. Immunol. 30, 111-112, 1993

A/Title: Comparison of complete nucleotide sequence of the human IgM heavy chain constant  
A/Reference number: S37767; MUID:93109369; PMID:8417370  
A/Accession: S37768

A/Molecule type: mRNA  
A/Residues: 1-453 <HAR>  
A/Cross-references: EMBL:X67301; NID:938407; PIDN:CAA77714.1; PID:938408  
A/Experimental source: cell line Ab 63  
C/Genetics:  
A/Map position: 14q32  
C/Superfamily: immunoglobulin C region; immunoglobulin homology

C/Keywords: glycoprotein; heterotrimer; immunoglobulin  
F/21-90/Domain: immunoglobulin homology <IMM1>  
F/127-199/Domain: immunoglobulin homology <IMM2>  
F/237-305/Domain: immunoglobulin homology <IMM3>  
F/344-415/Domain: immunoglobulin homology <IMM4>  
F/14/Disulfide bonds: interchain (to light chain) #status predicted  
F/28-88,134-197,244-303,351-413/Disulfide bonds: #status predicted  
F/46,209,272,279,440/Binding site: carbohydrate (asn) (covalent) #status predicted  
F/214,452/Disulfide bonds: interchain (to heavy chain) #status predicted  
F/291/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 14.2%; Score 385; DB 2; Length 453;  
Best Local Similarity 24.9%; Pred. No. 1.1e-15;  
Matches 114; Conservative 73; Mismatches 164; Indels 106; Gaps 15;

QY 36 TVELTCTASQ--KKSIOFHKNKSNQIKILGNQGSFLTGPBKNDRADSRSLMDQGNFP 93  
DB 23 SVAAGCLADPFLPDSITFSWKYKNNSDISSTRG-----FP 57  
QY 94 LIINKLIKIEDSDTYICEVED---QKEVQLVFGLTANSDTLHLOGOSLTLTLESPPGS 149  
DB 58 SVLRGKRYAATSOVLTPSKDVMQGTDEHVCVKQHPGNKKNVPLPVIAELPPKV 113  
QY 150 SPVQCR-----SPR-----GKNIOGKTLSTLSQLELQDSC- 180  
DB 114 SVFPPRDGFNGPRKSKLICQATGSPRQIOVSWLRGKQVSGVTTDQVQA 173  
QY 181 -----TWCTVLOKOKVEPKIDVPCAPPEKSCDKTTCPEL 220  
DB 174 TTYKVTSTLTIKESDMLSQSMFTCRV--DHKGLTFQONASSMCPVDDTAIR- 223  
QY 221 GGSVFLPFPKPDLMISRTPEVTCVVDVSHEDPEVKENWYDVGEVNAKPREEQ 280  
DB 224 -----VFALPPS-FASIFLTSTKTLCLVTLDTLTYD-SVTISMTQNGQAVKTH 277  
QY 281 YNSTYRVSVLTJLHODMLNGKEYKCKVSNKALPAPIEKTISSAKGQP-REPQ 339  
DB 278 PNAFTSAVGEASICEEDMNSGERFTCTVHTDLPSPKQITISRPKGVALLRPD 337  
QY 340 RDELT-KNOVSLTCLVKGFPSPDIIVAVESNGQP--ENNYKTTTPVLD--SDG 394  
DB 338 REQLNLRRESATITCLVTGFPADVFVOMQORGPLSEKVTYSAFMEPPAPG 397  
QY 395 LTVDKSRMOQGNVSCSVMEALHNHYTOKSLSPG 431  
DB 398 LTVSEEMNGETTYTCVAHAELPNRTERTVDSK 434

## RESULT 55

S25705  
Ig mu chain - sheep  
C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 23-Jul-1999  
C/Accession: S25705  
R/Facit, S.; Nau, F.  
Mol. Immunol. 29, 829-836, 1992  
A/Title: Isolation and sequence of a cDNA coding for the immunoglobulin mu chain of the  
A/Reference number: S25705; MUID:92342148; PMID:1635560  
A/Accession: S25705  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-592 <PAT>  
A/Cross-references: EMBL:X59994; NID:91269; PIDN:CAA42611.1; PID:912670

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:483-554/Domain: immunoglobulin homology <IMM>

Query Match 14.2%; Score 381; DB 2; Length 592;  
Best Local Similarity 27.0%; Pred. No. 1.8e-15;  
Matches 123; Conservative 78; Mismatches 153; Indels 102; Gaps 23;

```

QY 35 DTVELCTASQ--KKSIOFHKNSNOIKILNGSGSLTGKPSKLNRAISRSLMGOQNF 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 161 NYVALGCLADLPFNVSFSKFNSTVSEKRWTF---PEVLKD-----GLMSASSQ 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 93 PLIKKLIKEDSDTY-ICEVEDQK-----EEVQLL-----VEGLTANSDTLLQG 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 211 VALHSSTFOGTGYLVCEVQHPKEDVGHKVPRVEVLSPVSVFVPPCNS----- 263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 QSLTILTEPSPGSSPVQCR---SPR-----GKNI-----OQCKT--- 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 264 -----LSNGNSKSLICQATDPSPKQISLSWFRDKRIVSDISEGQVETVQSSPTTYR 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 169 -----LSVSOLELDQSGTWTCTVLQONOKVEFKIDIPCARPEPKSCDKHTPELLGSPS 224
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 318 AIVSLTITERMISGSAVYTCQVEHAKETQKAS-----SSCDATPPSP--IG--- 363
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 VFLPPEPKPDITMISRTPEVTCVVVDVSHEDPEVKENM-YVDG--VEVNAKTKPREEY 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 364 VETIPSPFAD-IFLTSAKLSCLVTNMLASYD--GLNISWHQNGKALETTTY----PERIL 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 282 NSTYRVASVLTVLHQDMLNGKEYKCKVSKNALPAPIEKTISKAG--QPREPOVYTLPPSR 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 NPTFSGRGRASVCSSEMGEEETCTVAHLDPFEKSAISKPKDAMKPSYVVLPPR 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 DELT-KNOVSLTCLVGFYPSDIAVESNGQP--ENNYKTPPVLD--SDGSFLYKSL 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 EQSLRESASVTLVGFAPADVFQVLOKGEPAASAKYVTSPPADEPDPSPAFVHSL 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 TVDKSRMOQGNVFSQVMHEALNHYTOKSLSPG 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 TVTEEDMSKERTYTCVGHGHALPHMWTERTVDKSTG 573
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 56

MHRUM

Ig mu chain C region, membrane-bound splice form - human

C:Species: Homo sapiens (man)

C>Date: 29-Jan-1993 #sequence\_revision 23-Aug-1997 #text\_change 22-Jun-1999

C:Accession: S16510; S09357; S16656; B26243; A02167

R:Dorai, H.; Gillies, S.D.

Nucleic Acids Res. 17, 6412, 1989

A>Title: The complete nucleotide sequence of a human immunoglobulin genomic C-mu gene.

A:Reference number: S09357; MUID:89366650; PMID:2505237

A:Accession: S16510

A:Molecule type: DNA

A:Residues: 433-473 <DOR1>

A:Cross-references: EMBL:X14939

A:Note: the authors translated the codon AAC for residue 445 as Met

A:Note: the sequence of residues 1-432 was assumed to be identical with the correspondin

A:Molecule type: DNA

A:Residues: 1-432, 'GKPTLVNVSIVMSDPTACTCY' <DOR2>

A:Cross-references: EMBL:X14940

A:Note: the authors translated the codon AAT for residue 16 as Met

A:Note: secreted splice form

R:Dorai, H.

submitted to the EMBL Data Library, April 1989

A:Reference number: S16656

A:Accession: S16656

A:Molecule type: DNA

A:Residues: 1-39, 'L', 41-432, 'GKPTLVNVSIVMSDPTACTCY' <DOR3>

A:Cross-references: EMBL:X14940

R:Rabbits, T.H.; Forester, A.; Milstein, C.P.

Nucleic Acids Res. 9, 4509-4524, 1981

A>Title: Human immunoglobulin heavy chain genes: evolutionary comparisons of C-mu, C-del

A:Reference number: A26243; MUID:82059479; PMID:6795593

A:Accession: B26243

A:Molecule type: DNA

A:Residues: 433-436, 'N', 438, 'E', 440-447, 'T', 449-473 <RAB>

A:Cross-references: GB:K01310; MUID:9184715; PIDN:AA859422.1; PID:9184734

C:Comment: During differentiation, B lymphocytes switch from expression of membrane-bou

A:Gene: GDB:IGHM

A:Cross-references: GDB:120086; OMIM:147020

A:Map position: 14q32.33-14q32.33

A:Intron: 1/1; 105/1; 217/1; 323/1; 433/1; 471/3

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kai

hain disulfide bonds. The IGM subunits associate into disulfide linked pentamers.

C:Superfamily: immunoglobulin C region; immunoglobulin homology <IMM1>

C:Keywords: alternative splicing; glycoprotein; heterotetramer; immunoglobulin; membran

F:21-90/Domain: immunoglobulin homology <IMM1>

F:127-199/Domain: immunoglobulin homology <IMM2>

F:237-305/Domain: immunoglobulin homology <IMM4>

F:344-415/Domain: immunoglobulin homology <IMM4>

F:433-473/Domain: carboxyl-terminal #status predicted <CTS>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:28-88,134-197,244-303,351-413/Disulfide bonds: #status experimental

F:46,209,272,279/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:214/Disulfide bonds: interchain (to heavy chain) #status experimental

F:291/Disulfide bonds: interchain (to mu chain in another subunit) #status experimental

Query Match 14.1%; Score 381; DB 1; Length 473;  
Best Local Similarity 23.9%; Pred. No. 2.1e-15;  
Matches 116; Conservative 77; Mismatches 166; Indels 126; Gaps 17;

```

QY 36 TVELCTASQ--KKSIOFHKNSNOIKILNGSGSLTGKPSKLNRAISRSLMGOQNF 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 23 SAVAGCLADLPDSITFSKTKNNSDISTRG-----PP 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94 LIKKLIKEDSDTY-ICEVEDQKEEVQLLVFGLTANSDTLL-----OQSILTYT 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 58 SVLRGKKAATSQVLLPSD-----VMQSTDEHVCKVQHPNKNKKNPLPIY 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 143 LESPPGSSPVQCR-----SPR-----GKNIQSGKTLVSQ 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 107 AELPPEKVSFVPPRGFGFNPRKSLICQATGSPROQLQVSWLRGKQVSGVTTDQYA 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 ELQDGG-----TWTCVLQONOKVEFKIDIVPEPAPEPSCDT 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 167 EKKEGPTTYKYSTLTIKESDWLQSMFTCV--DHGGLTFQONASSMCPDOPTAIR- 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 214 HTCPRLAGPSVFLPPEPKPDITMISRTPEVTCVVVDVSHEDPEVKFMVYDGVVHNAK 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 224 -----VFALPPS-FASIFLTSTKLTLCLVTLDTLTYD-SVTISWTRNGEAVKTN 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 274 TKPREQYNSTYRVASVLTVLHQDMLNGKEYKCKVSKNALPAPIEKTISKAGQP-REPO 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 271 TNISSEHNAPFSAVGEASICEDWNSEGRFCTVTHIDLPELQKTSIRPGVALHHRD 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 333 VYTLPPSDELDT-KNOVSLTCLVKGFFSYDIAVESNGQP--ENNYKTPPVLD--SDG 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 VYLLPAREQLNLRESATITCLVTGFSPADVFQVMQGGQPLSPKRYTASMPPPQAFG 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 388 SFFLYSKLTVDKSRMOQGNVFSQVMHEALNHYTOKSLSPGLQDLTCAEADQDELD 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 391 RYFASISILTVSEENMTGETTYC-VAHALLPRKWTERTVDNS-----TBGEVSADBEGR 444
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 GLWTT 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 445 NLWAT 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

## RESULT 57

S38864

Ig epsilon chain C region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 24-May-2001



F:46,209,272,279,439/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:214,451/Disulfide bonds: interchain (to mu chain in another subunit) #status experimental  
F:291/Disulfide bonds: interchain (to mu chain in another subunit) #status experimental

Query Match 13.9%; Score 376.5; DB 1; Length 452;  
Best Local Similarity 24.1%; Pred. No. 3.6e-15;

Matches 112; Conservative 75; Mismatches 156; Indels 121; Gaps 16;

```

QY 36 TVELTCTAG--KKSIQFHMKNSNQIKIIGNOSFLTTPSKLNDADSRSLMDGNRP 93
DB 23 SVAAGGLADFLPDSITFSMKYKNSDISTRG-----
QY 94 LIIKLIKIEDSDTYICEVEDOKEVOLLVFGLTANSDFHL-----QGQSLTPT 142
DB 58 SVLRGGKVAATSGVLLPSKD-----WQGTDEHVCKVQHPNKKENAVLPVI 106
QY 143 LSSPPGSPSVOCR-----SPR-----GKNIQGGKTLVSQI 174
DB 107 AELPFPKSVFVPRDGFQGNPKSKLICATGFPSPQIQVSWLRGKQVSGVTTDQVQA 166
QY 175 ELQDSG-----TWTCVLOQKQVFEKIDIVPCPAPRPKSCDT 213
DB 167 EKESGPTTYKVTSTLTIKESDMLGQSMFTCRV--DHRGLTFQNASMCVDPQDTAIR- 223
QY 214 HTCPBLGGSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAK 273
DB 224 -----VPAIRPS-FASIFLTKSKTKTCLTVDLTPTD-SVSISTRONGEAVKTH 270
QY 274 TYPREQVNTYRVVSVLTVLHODWLNGKEYCKVSNKALPAIEKTIKAKQOP-REPQ 332
DB 271 TWISESHPATFSAVEASICEEDMNSGERFTCTVHTDLPSPKQTIIRPKGVALLHRD 330
QY 333 VYTLPSRDELTKQVSLTCLVKGFPYBDIAVEMSNQOP--ENNYKTTPTPLD--SG 387
DB 331 VYLLPAPARQMLMRASATTCCTLTGSPADVFQVMQROQPSPEKTYVSAPWPEQAIQ 390
QY 388 SFPLYSKLTIVDKSRMOQGVNFSQVMEALHNHYTKSLSLSPG 431
DB 391 RFPANSHLVSEEMWTGETTTC-VAHEALPNRYTERIVYDKSG 433

```

# RESULT 59

S03186  
Ig heavy chain C region - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #ext\_change 20-Sep-1999

C:Accession: S03186  
R:Haite, R.N.; Shambloet, M.J.; Amemiya, C.T.; Litman, G.W.

Nucleic Acids Res. 17, 1776, 1989  
A:Title: A second Xenopus immunoglobulin heavy chain constant region isotype gene.

A:Reference number: S03186  
A:Accession: S03186

A:Molecule type: mRNA  
A:Residues: 1-448 <HA1>

A:Cross-references: EMBL:X13779; NID:G64827; PID:G64828  
A:Note: The authors translated the codon TTT for residue 9 as Ser and ATG for residue 34

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match 13.8%; Score 374; DB 2; Length 448;  
Best Local Similarity 26.8%; Pred. No. 5e-15;  
Matches 117; Conservative 57; Mismatches 144; Indels 118; Gaps 16;

```

QY 80 ADSRSLMDQNFPLIINKLK-IEDSD-----TYICEV---EDQ 114
DB 35 ADPLDIQMDGSLTITIKTRPVLSDVDGLYTLSQLTLASBWKSTYCKKVHNYTNT 94
QY 115 KEKVQLLVFGLTANSDFHLTQSGSLTLTESPPGSSP----- 151
DB 95 KQEKSLKVLPCMA-----PHYQLFLOSPCMSDAISRAGHENINATLDELCTINNY 145
QY 152 -----SVQCRSPRGKNIGGKTLVSQLELQDSGTW-----TCTVLQN 189

```

DB 146 HQGIKVKMLVNGKQDVSAEASVPTEPKTEDGTSSSSQLRII-KGMNKGTVSCIVTHT 204  
QY 190 QKKVEKIDIVCPAPRPSPCKHTCPBLGSPVFLFPKPKOTLMISRPETVCV- 248

DB 205 SSNTTTIANISQC-----TEQCHDNL--QVPLTFPHD-LYFSRNAKVTCLVS 250  
QY 249 -----VDVSHEDPEV-KFNMYVDGVEVHNAKTRERQVNTYRVVSVLTVLHODMNL 300

DB 251 SKKTIENPDISWREKAGLVEFTEDEPVLIHD-----NGTISVASILVCAEDWES 300  
QY 301 GKRYCKVSNKALPAIEKTIKSKA-KQPREQVNTYRVVSVLTLPSPRDELTKNQ-VSLTCLVKGFP 358

DB 301 GDFKSCVTSQDLPSPVKTIKFKONEGTPKADVDVLLPSSAQELIQEAWTLTCCFPTGPN 360  
QY 359 PSDIAVEMSNQ---QPENNYKTTPTPLVDSDFSFLYSGLTIVDKSRMOQGVNFSQVME 415

DB 361 PKEIFQMQGVNISSEDFINTVPKRSQEQTYFLYSKLAIPAKMNQGDVFTCVVGH 420  
QY 416 ALHNHYTKSLSLSPG 431

DB 421 ALPLVITQOSIDKSSG 436

# RESULT 60

MHMSM

Ig mu chain C region, membrane-bound form - mouse  
C:Species: Mus musculus (house mouse)

C:Date: 31-Oct-1980 #sequence\_revision 30-Jun-1991 #ext\_change 22-Jun-1999  
C:Accession: A02167; A37517; B02166

R:Barly, P.; Rogers, J.; Davis, M.; Calame, K.; Bond, M.; Wall, R.; Hood, L.  
Cell 20, 313-319, 1980

A:Title: Two mRNAs can be produced from a single immunoglobulin mu gene by alternative  
A:Reference number: A02167; MUID:80222874; PMID:6771020

A:Accession: A02167  
A:Molecule type: DNA

A:Residues: 436-476 <EAR>  
A:Cross-references: GB:J00816; GB:J00444; NID:G52343; PID:CAA24197.1; PID:G52344

R:Rogers, J.; Barly, P.; Carter, C.; Calame, K.; Bond, M.; Hood, L.; Wall, R.  
Cell 20, 303-312, 1980

A:Title: Two mRNAs with different 3' ends encode membrane-bound and secreted forms of  
A:Reference number: A37517; MUID:80222873; PMID:6771019

A:Accession: A37517  
A:Molecule type: mRNA

A:Residues: 410-476 <ROG>  
A:Cross-references: GB:V00821; NID:G52355; PID:CAA24202.1; PID:G817972

R:Kawakami, T.; Takahashi, N.; Honjo, T.  
Nucleic Acids Res. 8, 3933-3945, 1980

A:Title: Complete nucleotide sequence of mouse immunoglobulin mu gene and comparison with  
A:Reference number: A02166; MUID:81076590; PMID:6255422

A:Accession: B02166  
A:Molecule type: DNA

A:Residues: 1-435; GKPLVNVSHMSPTGTCV- <KAW>  
A:Comment: The sequence of residues 1-409 was assumed to be identical with the correspond

C:Comment: During differentiation, B lymphocytes switch from expression of membrane-bound  
S:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (lambda)

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin

F:21-91/Domain: immunoglobulin homology <IMM1>  
F:129-201/Domain: immunoglobulin homology <IMM2>

F:239-307/Domain: immunoglobulin homology <IMM3>  
F:346-417/Domain: immunoglobulin homology <IMM4>

F:436-476/Domain: carboxyl-terminal <CTS>  
F:14/Disulfide bonds: interchain (to light chain) #status predicted

F:28-89,136-199,246-305,353-415/Disulfide bonds: #status predicted  
F:46,211,243,258,281/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:16/Disulfide bonds: interchain (to heavy chain) #status predicted  
F:293/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.8%; Score 373.5; DB 1; Length 476;  
Best Local Similarity 25.2%; Pred. No. 5.8e-15;

```

QY 421 ALPLVITQOSIDKSSG 436

```







```

Db      208 FEDKAVMSSECTFSP-----GIQVPIAPSFADT-FLSKSARLICTL 250
Qy      248 VVDVSHDEDEVKFNVVDGVEVHNAKT-----KPREQYNSTRYRVSVLTVLHODMLNCK 302
Db      251 VYDLTLYTG-SLNI-SW-----ASHNGALDTHMNITESHNPATFSAMGASVCAEMESGE 304
Qy      303 EYKCKVSNKALPAPIEKTISKAKGPRE-POVYTLPPSRDEL-TKNQVSLTCLVKGFPYS 360
Db      305 QCTCTVTHADLPFLKHTISKSEVAKHPAYVLPAPAEQLVRESATVTCLVKGFSFA 364
Qy      361 DIAVESNGQP-ENNYKTTPVLD--SDGSFELYSKLTVDSRMQGVNFCSCVMEHA 416
Db      365 DVFVGQORQGPISDPKYVTSAPAPAPQAPGLYFTHTSLTLTVEEDMNSGETFTCVVGHEA 424
Qy      417 LNMHTQKSLSPGLQDTECAEADGELDGLMTT 452
Db      425 LPHMVTERTVDS-----TEGEVGAEEEGFENLMTT 455

```

## RESULT 63

1g mu chain C region - dog (tentative sequence)

C/Species: Canis lupus familiaris (dog)  
 C/Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 31-Mar-2000  
 C/Accession: A93131; A94246; A02169  
 R/McCumber, L.J.; Capra, J.D.  
 Mol. Immunol. 16, 565-570, 1979  
 A/Title: The complete amino-acid sequence of a canine mu chain.  
 A/Reference number: A93131; MUID:80077682; PMID:117299  
 A/Contents: myeloma protein Moo  
 A/Accession: A93131  
 A/Molecule type: protein  
 A/Residues: 1-177 <MCC>  
 R/Wasserman, R.L.; Capra, J.D.  
 Science 200, 1159-1161, 1978  
 A/Title: Amino acid sequence of the Cc region of a canine immunoglobulin M: interspecies  
 A/Reference number: A94246; MUID:78180587; PMID:653360  
 A/Contents: Moo

A/Accession: A94246  
 A/Molecule type: protein  
 A/Residues: 178-450 <WAS>  
 C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) heavy disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C/Superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F/20-89/Domain: immunoglobulin homology <IMM1>  
 F/126-196/Domain: immunoglobulin homology <IMM2>  
 F/234-302/Domain: immunoglobulin homology <IMM3>  
 F/341-413/Domain: immunoglobulin homology <IMM4>  
 F/430-450/Domain: immunoglobulin homology <CTS>  
 F/14/D/Disulfide bonds: interchain (to light chain) #status predicted  
 F/27-87,133-194,243-300,348-410/Disulfide bonds: #status predicted  
 F/50,206,269,276,437/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F/211,449/Disulfide bonds: interchain (to heavy chain) #status predicted  
 F/288/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.7%; Score 369; DB 1; Length 450;  
 Best Local Similarity 24.1%; Pred. No. 1e-14;  
 Matches 119; Conservative 80; Mismatches 139; Indels 156; Gaps 20;

```

Qy      36 TVELTCTAG--KKSQIFHWKNSNQIKLNGGSLFKTPSKLNDADSRSLMDGNRP 93
Db      22 TVAMGGLADFLPGSLTFMSKYEBSAISNSTRG-----PP 56
Qy      94 LIIKMLK-----IEDSPTI-CEVE-----DOKEVOLVFGLTANSPTHL 134
Db      57 SVLRGKYATISQVFLPSVDIIQGTDEHIVCKVRHSBGRKQKQKVPVFN----- 105
Qy      135 QGOSLTLLSPGSSPSVOCR-----SPR-----GKNIQGGKT 168
Db      106 -----LTL--PPEVSGFIPPRDAFPGBRKSQLICQAGSFSPRQVSLRDGQIISGVT 157

```

```

Qy      169 LSVQLEIODSG-----TWTC-----TVLQNKKEFKIDIV 200
Db      158 TNEVAZAKZSGPTTYKVTSMULTIOEDAMLSGVFTCVKEHNGLTFQGNASSM----- 210
Qy      201 PCPAPKPSCDKTHTCPELLGGPSVFLPPPKKDTLMTISRPEVTCVVDVSHDEPEYK 260
Db      211 -CTSDQPV-----GISIFTIIPS-FASIFMTKSAKLCLVTLDTLATVY-SVTI 254
Qy      261 NMVYDGEVHNAKTPREQYNSTRYRVSVLTVLHODMLNGEYKCKVSNKALPAPIEKT 320
Db      255 SWTRENGALKTHNIISSHPNGTFSAMGEATVCEVESGQFTCTVHTDLPVSLKOT 314
Qy      321 ISKARG-OPREPOVYTLPPSRDEL-TKNQVSLTCLVKGFPSPDIAMVESNGQ--PENNY 376
Db      315 ISRPKGVAVHNPSSVVLPPSRQQLDRESATLSCLVTGSPSPDVFQWQKQPVPPDSY 374
Qy      377 KTTTPVLD--SDGSFELYSKLTVDSRMQGVNFCSCVMEHAHNYTKSLSPG--- 431
Db      375 VTSAPMPPEQAPGLYFAHSILTVSEEMNAGETVGVVAHESLPNRVTERSVDSKGTGKPT 434
Qy      432 -----LQDETCAE 440
Db      435 LVNVSLLVSDTAGZ 448

```

## RESULT 64

1g mu heavy chain disease protein (Bot) - human

C/Species: Homo sapiens (man)  
 C/Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 16-Jul-1999  
 C/Accession: A02163  
 R/Barnikol-Matanabe, S.; Mihaesco, E.; Mihaesco, C.; Barnikol, H.U.; Hilschmann, N.  
 Hoppe-Seyler's Z. Physiol. Chem. 365, 105-118, 1984  
 A/Title: The primary structure of mu-chain-disease protein BOT. Peculiar amino-acid seq  
 A/Reference number: A02163; MUID:84184186; PMID:6425189  
 A/Accession: A02163  
 A/Molecule type: protein  
 A/Residues: 1-391 <BAR>  
 C/Comment: This protein has no V region homology or CH1 region.  
 C/Genetics:  
 A/Gene: GDB:IGHM  
 A/Cross-references: GDB:120086; OMIM:147020  
 A/Map position: 14q32.33-14q32.33  
 C/Superfamily: immunoglobulin C region; immunoglobulin homology  
 C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin; transmembrane pr  
 F/1-41/Domain: pre-C <VAR>  
 F/43-391/Domain: Ig mu chain C region, secreted form <IGM>  
 F/65-137/Domain: immunoglobulin homology <IMM1>  
 F/175-243/Domain: immunoglobulin homology <IMM2>  
 F/282-353/Domain: immunoglobulin homology <IMM3>  
 F/147,210,217,378/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.5%; Score 365.5; DB 1; Length 391;  
 Best Local Similarity 28.0%; Pred. No. 1.4e-14;  
 Matches 105; Conservative 65; Mismatches 132; Indels 73; Gaps 15;

```

Qy      110 EYED---QKEVQVLVFGLTANSPTHLIQQGSLTFLSPGSSPSVOCR----- 156
Db      18 EAEDRIIRKEEARL-----SGRD---MQVSIQPIAELPPKVSFVPPRDSFGPNRKS 68
Qy      157 -----SPR-----GKNIQGGKTLVSQLEIODSG-----TWTCVLYON--- 189
Db      69 KLICQATFSRQIIVSWLRBSQVGSVTTDEVAEAKESGPTTYKVTSTLTIKESDML 128
Qy      150 -QKVEFKIDVPCPAPKPSCDKTHTCPELLGGPS-----VFLPPPKKDTLMTISRTP 242
Db      129 QGSMTCVDRHNGLTFQGNASS-----MCGSDQDTAIRVFAIPPS-FASIFLTST 178
Qy      243 EYTCVAVVDVSHDEPEVKFNWYDVGEVHNAKTPREQYNSTRYRVSVLTVLHODMLNCK 302
Db      179 KLTCLVTLTLTYD-SVTISWTRQDEAVKTHNIISSHPNATFSAVGASLIEDDWDSDGE 237
Qy      303 EYKCKVSNKALPAPIEKTISKAKGP-REPQVYTLPPSRDEL-TKNQVSLTCLVKGFPYS 360

```

```

Db      238 RFTCTVHTDLPSPKQITSRPKGVALHARPVYLLPAPAREQLINRESATITCLVTGFGSPA 297
Oy      361 DIAWESNGOP--ENNKTTPVVD--SDGSFPLXSLTVDKSMOQGNFSCSVMEHA 416
Db      298 DVFVOMRGQDLSEKVTYSAWPEPOAPKRYFAHSLITVSEEMNGETTYTCAVAHEA 357
Oy      417 LHNHYTQKLSLSPG 431
Db      358 LPNRYTERTVDKSTG 372

```

## RESULT 65

EMMS

Ig epsilon chain C region (version 2) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Aug-1996

C:Accession: A02145

R:Title: N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.

EMBO J. 1, 1117-1123, 1982

A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison w/

A:Reference number: A90966; MUID:84236092; PMID:6329728

A:Accession: A02145

A:Molecule type: DNA

A:Residues: 1-423 &lt;ISH&gt;

A&gt;Note: the sequence was determined from the germline gene

C:Genetics:

A:Introns: 91/1; 199/1; 307/1  
 C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 16 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin  
 F:16-77/Domain: immunoglobulin homology <IMM1>  
 F:115-183/Domain: immunoglobulin homology <IMM2>  
 F:220-288/Domain: immunoglobulin homology <IMM3>  
 F:325-396/Domain: immunoglobulin homology <IMM4>  
 F:123-75,122-181,227-286,332-394/Disulfide bonds: #status predicted  
 F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.5%; Score 363.5; DB 1; Length 423;

Best Local Similarity 29.8%; Pred. No. 2e-14; Indels 61; Gaps 17;

Matches 103; Conservative 68; Mismatches 114;

```

Oy      109 CEVEDQKEEVDL--LVFGLTAN--SDTHLLOGSLTTLTSPGSSPVSQCRSPRKNIQ 164
Db      109 CDPAHFSTIDLYGCTYGHINDVSVMIMDRBTDL-----ACQVILKE--EGKLAS 161
Oy      165 GSKTISVSQLELDGSGTCTVVLQNOQKVEFKIDIVCPAPRPSCKDHTTCPELLGSPS 224
Db      162 TCSKLTINTEQOMWSESTFCRV--TSQGVLYLAHTRRCPDHEPR-----GAI 206
Oy      225 VFLPPPKDITLMTSRPEVTCVVVDV--SHEDPEVKFN-----WYVDGVEVYNA 272
Db      207 TYLLPSPSLD-LYQAGAPKLTCLVVDLSEKNVNTWMOEKTSVSASQWY--TKHN- 261
Oy      273 KTKPREQNSTRYVSVLYTLVHODPLNGEYKCKVSKALPAIEKTSISKAK--GQPREP 331
Db      262 -----NATSTISLTPVAAKQIMRGYQGVCDVDPDPKPRIVRSITILPQVSQSRAP 312
Oy      332 QVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQDENN--YKTPPVLDSDGS- 388
Db      313 EYVYFPPPEEE--SEKRTLTCLIQVFPEDISVQWLGDKLISNSQHSHTTP--LKSNGSN 370
Oy      389 --FFLYSKLTVDKSRMOQGNFSCSVMEALHN--HYTKSLISLSPG 431
Db      371 QGFPLFSKLEAVAKTLMTQKQPTCOVHIALQKPRKLEKTIISTISLG 416

```

## RESULT 66

A24976

Ig mu chain C region (allele b) - mouse

C:Species: Mus musculus (house mouse)

C:Date: 05-Jun-1988 #sequence\_revision 05-Jun-1988 #text\_change 21-Jan-2000

C:Accession: A24976  
 R:Schröter, P.H.; Quester, S.; Bothwell, A.  
 Nucleic Acids Res. 14, 2381-2389, 1986  
 A:Title: Alleotypic differences in murine mu-genes.  
 A:Reference number: A24976; MUID:86176735; PMID:3083402  
 A:Accession: A24976  
 A:Molecule type: mRNA  
 A:Residues: 1-455 <SCH>  
 A:Cross-references: GB:X03690; NID:952381; PIDN:CAA27326.1; PID:952382  
 A:Experimental source: strain C57BL/6  
 A>Note: the authors translated the codon AAG for residue 65 as Leu  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:346-417/Domain: immunoglobulin homology <IMM>

Query Match 13.4%; Score 363; DB 2; Length 455;

Best Local Similarity 25.3%; Pred. No. 2.3e-14; Indels 110; Gaps 22;

Matches 116; Conservative 84; Mismatches 149;

```

Oy      37 VELTCTASQ--KKSIOFHKNSQIKILGNGSFPLTKGPKSLANDRARSRLMDQGNF-- 92
Db      24 VAMGCLARDLPSTISTYNYQNNTVEYIOGIRTPPT-----LRTGKLYLA 68
Oy      93 ---PLIKNLKIEDSDTY--ICEVEDQKEEVQLVFGTLTANSDDL-----LQGSLLTLE 144
Db      69 TSQVLSPKSILEGSDDEVLVCKIH-----YG--GKNKDLHVPIPAVLENNPNVNVF 117
Oy      145 SPR-----GSSP---SVQCR-----SPR-----GKNIOG----- 166
Db      118 VPPRDFGSGAPRPSKILCATATNTPPRITVSWLKDGLVESGTTDPVTIENKSPPTQT 177
Oy      167 ----KTLVSQLELDGSGTCTVVLQNOQKVEFKIDIVCPAPRPSCKDHTTCPELLG 222
Db      178 YKVISTILTEIDMLNANVTCTV--DHRGLTFLKAVNSITCAASPT-----DIL-- 225
Oy      223 PSVFLPPPKDITLMTSRPEVTCVVVDVSHEDPEVKFNMYVDGVEYNAKTPREQYN 282
Db      226 --FTTIPSPFAD--IFLSKSNLITCLVSNLTYE--TLNISWASQGEPLKTIKIMESH 281
Oy      283 STYVSVVLVLYHODPLNGEYKCKVSKALPAIEKTSISKAKQPRE-----QVYTLR 337
Db      282 GTFSAKVAASVCEVDMMNRKEFVCTVTRDLPSPQKFKISK---PNEVHGRPAVYLLP 337
Oy      338 PSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQ--PENNYKTPPVLD--SDGSFPLY 392
Db      338 PARQLMRSATYTCIVKGFSPADISVQWLGQQLLPQEKTYTSAWPEGAGCFYFTH 397
Oy      393 SKLTVDKSRMOQGNFSCSVMEALHNHYTQKSLISLSPG 431
Db      398 SILTVTEENWNGEYTCVGVGHEALPHLVTERVTVDKSTG 436

```

## RESULT 67

MEMS

Ig mu chain C region, secreted form - mouse

C:Species: Mus musculus (house mouse)

C:Date: 31-Oct-1980 #sequence\_revision 31-Oct-1980 #text\_change 16-Aug-1996

C:Accession: A02166; A26239; A26240; B02039

R:Kawakami, T.; Takahashi, N.; Honjo, T.

Nucleic Acids Res. 8, 3933-3945, 1980

A:Title: Complete nucleotide sequence of mouse immunoglobulin mu gene and comparison with

A:Reference number: A02166; MUID:81076590; PMID:6255422

A:Accession: A02166

A:Molecule type: DNA

A:Residues: 1-455 &lt;KAW&gt;

A:Cross-references: GB:J00443

A&gt;Note: the sequence was determined from the germline gene

R:Goldberg, G.I.; Vanin, E.F.; Zrolka, A.M.; Blattner, F.R.

Gene 15, 33-42, 1981

A:Title: Sequence of the gene for the constant region of the mu chain of Balb/c mouse imm

A:Reference number: A26239; MUID:82051295; PMID:6795090

A:Accession: A26239

A:Molecule type: DNA

A:Residues: 1-455 <GOL>  
 A:Note: the sequence was determined from the germ-line gene  
 R:Auffray, C.; Rougeon, F.  
 Gene 12, 77-86, 1980  
 A>Title: Nucleotide sequence of a cloned cDNA corresponding to secreted mu chain of mouse  
 A:Reference number: A26240; MUID:81165562; PMID:6260591  
 A:Accession: TEPC183  
 A:Contest: A26240  
 A:Accession: A26240  
 A:Molecule type: mRNA  
 A:Residues: 1-225, 'N', 227-257, 'S', 259-367, 'K', 369-455 <NUP>  
 R:Kenly, M.; Sibley, C.; Fuhrman, J.; Schilling, J.; Hood, L.E.  
 Proc. Natl. Acad. Sci. U.S.A. 76, 2932-2936, 1979  
 A:Reference number: A26241; MUID:79223904; PMID:111247  
 A:Contest: annotation; MOPC 104E  
 A:Note: this sequence has been revised in reference A02039. Carbohydrate binding sites a  
 R:Kenly, M.R.; Fuhrman, J.S.; Schilling, J.W.; Rogers, J.; Sibley, C.H.; Hood, L.E.  
 Biochemistry 21, 5415-5424, 1982  
 A>Title: Complete amino acid sequence of a mouse mu chain: homology among heavy chain co  
 A:Reference number: A02039; MUID:83075344; PMID:6816276  
 A:Contest: MOPC 104E  
 A:Accession: B02039  
 A:Molecule type: protein  
 A:Residues: 1-77, 'N', 79-100, 'Q', 102-225, 'N', 227-257, 'T', 259-367, 'K', 369-455 <KEH>  
 C:Genetics:  
 A:Introns: 1/1; 106/1; 219/1; 325/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
 hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1a  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob  
 F:129-91/Domain: immunoglobulin homology <IMM1>  
 F:129-201/Domain: immunoglobulin homology <IMM2>  
 F:239-307/Domain: immunoglobulin homology <IMM3>  
 F:346-417/Domain: immunoglobulin homology <IMM4>  
 F:436-455/Domain: carboxyl-terminal <CTRS>  
 F:114/Disulfide bonds: interchain (to light chain) #status predicted  
 F:128-89/Disulfide bonds: #status experimental  
 F:46, 211, 243, 281, 442/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:136-199, 246-305, 353-415/Disulfide bonds: #status predicted  
 F:216 455/Disulfide bonds: interchain (to heavy chain) #status predicted  
 F:293/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.4%; Score 362; DB 1; Length 455;  
 Beat Local Similarity 25.3%; Pred. No. 2.6e-14;  
 Matches 116; Conservative 84; Mismatches 149; Indels 110; Gaps 22;

QY 37 VELTCTASQ--KKSIOFHKNKSNQIKILNQGSPFLTKGFSKLNDRADSRHSIMDQNF-- 92  
 DB 24 VAMGCLARDPLRSTISFTWNYONNTEVIOGIRTPPT-----LRTGAKYLA 68  
 QY 93 ---PLIKLKIKIEDSDTY-ICEVEDQKEVQLVGLTANSPTHL-----LOGQSILTLLE 144  
 DB 69 TSQVLLSPKSLIEGSDVLYCKIH-----YG-GKNRDLHVPITPAVAMENNVNVF 117  
 QY 145 SPP-----GSSP---SVQCR---SPR-----GKNIOGG----- 166  
 DB 118 VPRROGFGSGPARPKSKLICEATNFTPKPTTWSMLKDGKLVESGFTTDPVTIENKSTPOT 177  
 QY 167 ---KTLVSQLELDSDGTCTCTVLONQKVEFKIDIVCPAPRPSKCDKTHTCBELLGG 222  
 DB 178 YKVIKTLTITSEIDMINLVYTCRV--DHRGLTFELKNVSTCAASPST-----DIL-- 225  
 QY 223 PSVFLPFRPKDITMISRTPEVTCVVVDVSHEDPEVKFKNVYDGVENNAKTKPRREDOYN 282  
 DB 226 --TFTTPSPFAD-IFLISKANLITCLVSNLATE-TLNTSWASQSGEPLETIKIKIMESHEN 281  
 QY 283 STYRVSVTLVHODMLNKEKCKVSNKALPAPIEKTISKAKGPR-----POVYTLR 337  
 DB 282 GTFSAKGVASVCEVDENNRKEFVCTYTHDLSPQKPKTSK---PNEVHKRPRAVYILR 337  
 QY 338 PSRDEL-T-KNQVSLTCLVKGFPYSDIAVWESNGQ--PENNYKTPPVLD--SDGSPFLY 392  
 DB 338 PAREQNLNRESATVCTLVGFGSPADISVQWLQGLLPEKQVVTAPMPDEGAPGFYFTH 397

QY 393 SKLTVDKSRMOQGNVFCSVNHEALHNHYTKSISLSPG 431  
 DB 398 SILVTEENWNSGETYTCVVGHEALPHLVTERTVDXKSG 436

RESULT 68  
 136948  
 Ig epsilon-chain - chimpanzee (fragment)  
 C:Species: Pan troglodytes (chimpanzee)  
 C:Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 21-Jan-2000  
 C:Accession: 136948  
 R:Sakoyama, Y.; Hong, K.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987  
 A>Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and oranguta  
 A:Reference number: 136948; MUID:87147196; PMID:3103123  
 A:Accession: 136948  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-426 <RES>  
 A:Cross-references: GB:M15398; NID:g176797; PIDN:AA35416.1; PID:g176798  
 C:Genetics:  
 A:Introns: 103/1; 209/1; 317/1  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 13.4%; Score 361.5; DB 2; Length 426;  
 Beat Local Similarity 28.9%; Pred. No. 2.6e-14;  
 Matches 102; Conservative 54; Mismatches 136; Indels 61; Gaps 10;

QY 90 GNFPLIKLKIEDSDTYICEVEDQKEVQLVGLTAN--SDTHLOGQSILTLTBSP 147  
 DB 123 GHFPPTQLL-----CLVSGYTFGTINITWLEBQGVMDVDL----- 158  
 QY 148 GSSPSVQCRSPRGKNIQCGKTLVSQLELDSDGTCTCTVLONQKVEKIDIVCPAPR 207  
 DB 159 ---STASATBEGELASTQSEILTSQKMWLSDRYTCQVTVQGGTFE----- 201  
 QY 208 KSCDKTHCPRL-LGSPVFLPFRPKDITMISRTPEVTCVVVDVSHEDPEVKFKNVYD 266  
 DB 202 --DSTKCAQSNPRGSAVLSRSPFP-LTRISPTTCLVVDLAPSKGTVNLTWSPAS 257  
 QY 267 VEVNNAKTPREDOYNSTYRVSVTLVHODMLNKEKCKVSNKALPAPIEKTISKAKG 326  
 DB 258 GKPVNHSTRKEQKQNGLTITSTLPVGTROMIBETVQCVTHNPLRALVRSITKSG 317  
 QY 327 QPREPOVYTL-----PSRDELTKNQVSLTCLVKGFPYSDIAVWESN--GQPENNYKT 379  
 DB 318 PRAAPEVYAFATPREGSGRDKRT-----LACLIQWFPEDISVQWLHNEVOLPDARHSTT 372  
 QY 380 PVLVDSDSFFLYSLKTLTKDKSRMOQGNVFCSVNHEALHNHYT--OKSISLSPG 431  
 DB 373 QPHKTKSGGFVFSRLVETRAEWQKDFICRAVHAASPSQTVQRTVSNPG 425

RESULT 69  
 EHHU  
 Ig epsilon chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Mar-1981 #sequence\_revision 13-Jun-1983 #text\_change 16-Jul-1999  
 C:Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116; C46  
 R:Flanagan, J.G.; Rabbits, T.H.  
 EMBO J. 1, 655-660, 1982  
 A>Title: The sequence of a human immunoglobulin epsilon chain constant region gene  
 A:Reference number: A22771; MUID:84236029; PMID:6234164  
 A:Accession: A22771  
 A:Molecule type: DNA  
 A:Residues: 1-428 <FLA>  
 A:Cross-references: GB:L00022; GB:T00227; GB:V00555; NID:g185035  
 R:Ueda, S.; Nakai, S.; Nishida, Y.; Hasejima, H.; Honjo, T.  
 EMBO J. 1, 1539-1544, 1982  
 A>Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudog  
 A:Reference number: A23195; MUID:84207910; PMID:6327276  
 A:Accession: A23195

A:Molecule type: DNA  
A:Residues: 2-428 <UND>  
A:Cross-references: GB:J00222; NID:g184755  
R:Zhang, K.; Saxon, A.; Max, E.E.  
J. Exp. Med. 176, 233-243, 1992  
A>Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing  
A:Reference number: PH1214; MUID:92308839; PMID:1613458  
A:Accession: PH1214  
A:Molecule type: DNA  
A:Residues: 320-428 <ZHA>  
A:Cross-references: EMBL:X63693; GB:S38668; NID:g32987  
R:Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Kikuchi, M.; Sugita,  
Nucleic Acids Res. 11, 719-726, 1983  
A>Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon chain  
A:Reference number: A93491; MUID:83168897; PMID:6300763  
A:Accession: A93491  
A:Molecule type: mRNA  
A:Residues: 1-428 <SEN>  
A:Cross-references: GB:L00022; GB:U00227; GB:V00555; NID:g185035  
R:Max, E.E.; Battey, J.; Ney, R.; Kirsch, I.R.; Leder, P.  
Cell 29, 691-699, 1982  
A>Title: Duplication and deletion in the human immunoglobulin epsilon genes.  
A:Reference number: A90824; MUID:83001945; PMID:6288268  
A:Accession: A90824  
A:Molecule type: DNA  
A:Residues: 1-358, 'L', 360-428 <MAX>  
A:Cross-references: GB:U00222; NID:g184755  
A>Note: This sequence difference may be due to polymorphism  
R:Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.  
in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-3  
A:Reference number: A94418  
A:Accession: A94418  
A:Molecule type: Protein  
A:Residues: 'GMWL', 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 124  
A:Experimental source: myeloma protein Nd  
R:Kerten, J.H.; Molgaard, H.V.; Houghton, M.; Darbyshire, R.B.; Vinay, J.; Bell, L.O.; G  
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982  
A>Title: Cloning and sequence determination of the gene for the human immunoglobulin epsilon  
A:Reference number: A93933; MUID:83065234; PMID:6815656  
A:Accession: B93933  
A:Molecule type: mRNA  
A:Residues: 1-40:68-114:427-428 <KEN>  
A:Cross-references: GB:L00022; NID:g185035  
R:Ikeyama, S.  
FEBS Lett. 224, 306-310, 1987  
A>Title: Purification and characterization of a recombinant human IgE Fc-epsilon fragment  
A:Reference number: S02438; MUID:88083554; PMID:3121387  
A:Accession: S02438  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 98-352 <IKE>  
R:Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.  
J. Biol. Chem. 269, 456-462, 1994  
A>Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces  
A:Reference number: A53116; MUID:94105254; PMID:8276835  
A:Accession: A53116  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 320-428 <ZH2>  
A:Experimental source: myeloma U266-derived cell line AF-10  
A>Note: Sequence extracted from NCBI Backbone (NCBIN:141701, NCBIP:141702)  
R:Hellman, L.  
Eur. J. Immunol. 23, 159-167, 1993  
A>Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of  
A:Reference number: A46536; MUID:93122085; PMID:8419166  
A:Accession: C46536  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 382-426 <HEL>  
A:Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167  
A:Experimental source: B cell myeloma U-266  
A>Note: Sequence extracted from NCBI Backbone (NCBIP:125297)  
A:Accession: D46536

```

A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-391 <HE2>
A:Cross-references: GB:S55276; NID:Q263168; PIDN:AAB24858.1; PID:G263169
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:125299)
A:Accession: A46516
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 401-428 <HE3>
A:Cross-references: GB:S53497; NID:Q263162; PIDN:AAB24855.1; PID:G263163
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:123483)
C:Genetics:
A:Gene: GDB:IGHB
A:Cross-references: GDB:119335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterodimer subunit consists of two identical light (kap)
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterodimer; immunoglobu
F:12-87/Domain: immunoglobulin homology <IM1>
F:128-195/Domain: immunoglobulin homology <IM2>
F:232-301/Domain: immunoglobulin homology <IM3>
F:338-407/Domain: immunoglobulin homology <IM4>
F:134/Disulfide bonds: interchain (to light chain) #status predicted
F:15.105-29-85.135-299.345-405/Disulfide bonds: #status predicted
F:121.409.99.146.252.275/Bridging site: carbohydrate (Asn) #status experimental
F:121.209/Disulfide bonds: interchain (to heavy chain) #status predicted

Query Match 13.4%; Score 361.5; DB 1; Length 428;
Best Local Similarity 28.9%; Pred. No. 2.6e-14;
Matches 102; Conservative 53; Mismatches 137; Indels 61; Gaps 10;

QY 90 GNFPLIKNLKIEDSDTYICEVEDQKEEYQLVFGILTAN--SDTHLQGGSLTLTLESPP 147
DB 125 GHFPTLTQL-----CLVSGTTPGTINTLTWLDGQVMDVL----- 160
QY 148 GSSPVSQCRSPRGKNIQGGKTLVSQLELDSDSGTWCTTVLQNGKVEFKIDIVCPADPP 207
DB 161 ----STASTQGGELASTQSELTLSQKMLSDTYTCQVYVQGHF----- 203
QY 208 KSCDKHTTCPEL-LGSPSYFLPPPKPKDTLMTSRTEBTVGVVDVSHEDPEVKENMYVDG 266
DB 204 ---DSTKKCADSNPRGVSAYLSRSPSPD-LFIRKSPITCLVLDLAPSKGTVNTWSPAS 259
QY 267 VEVNHAKTKPREQVNSTRYVSVLTVLHODMLNGEKYKKKNSKALPAPREKITSKRG 326
DB 260 GKPVNHSTRKEKQKRNGLTITVSTLPGTRDWLEGTEYQQRVTHPHLPRALMRSTTKSG 319
QY 327 QPREQVYTL-----PPSRDELTKQNVSLTCLVKGFPSPDIAYEWESN--GQPNENYKTT 379
DB 320 PRAAEVVAFAFPMPGSDKRT----LACLQNWPEIDISQVMLNVEQLPARHSTT 374
QY 380 PPVLSDSDSFFLYSKLTYDKSRWQGNVFSQVMEALNHYT-QKSLSPG 431
DB 375 QPKTKGSGFFVFSRLLEVTRAEMEQKDFICRAVHEAASPQTVQRAVSVVPG 427

RESULT 70
S00390
I: gamma chain (clone 36) - chicken (fragment)
N: Alternate names: Ig nu chain
C: Species: Gallus gallus (chicken)
C: Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C: Accession: S00390
EMBO J. 7, 739-744, 1988
A: Title: Chicken immunoglobulin gamma-heavy chains: limited VH gene repertoire, combinat
A: Reference number: S00390; MUID:88283642; PMID:3135182
A: Accession: S00390
A: Molecule type: mRNA

```

A:Residues: 1-504 <PAR>  
 A:Cross-references: EMBL:X07174  
 A:Note: this sequence was determined from the differentiated gene  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin

Query Match 13.4%; Score 361; DB 2; Length 504;  
 Best Local Similarity 28.3%; Pred. No. 3.4e-14;  
 Matches 95; Conservative 70; Mismatches 129; Indels 42; Gaps 14;

QY 114 QKEEVL--LVFGLT-ANDTH-LLOG-OSLTLESPPGSSPSVOCRRPKNIQGGKT 168  
 DB 192 QSEVELLCVYGFSPASAEVEMLVGGLVVASQSPVRSSTYSSSR----- 242  
 QY 169 LSVSLELDGSGMTCTVQONOKKVEFKIDIVCPAPRPPKSCDKHTCPBELLGGSPVLF 228  
 DB 243 VAVSGTDMREGKSYSCRVAHPATNTVVEHVGKCP-----DGAQSCSPI---QLVVI 291  
 QY 229 PRPKDTLMISRTPEVTCVVVDVSHEDPEVKENMYVDGVEVHNAKTKPR--EQYNSTYR 286  
 DB 292 PPSGPE-LYISLDATKRLCVLNP--SDSSLVTWIRE--KSGNLRPDPVLDHFNIGTYS 347  
 QY 287 VVSULTVLHODMLNGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN 346  
 DB 348 ASSAVPVSTQDMLSGERFCTVQHEELPLPLSKSVYRNTGPTTPPLIYFPAPHPBELSLIS 407  
 QY 347 QVSLTCLVYGFYPSDIAVEM--ESNQGPENNYKTPPVLD-----SDG-SFLYSLK 395  
 DB 408 RYTLSCLVNFGFRPRDEIRMLDHRVAPTEFVYTTAVLPEERTANGAGDGTFFVYSKM 467  
 QY 396 TVDKSRMOQGNFVSCSVMEALHNHYTKSLSPG 431  
 DB 468 SVETAKMNGTVPACAVAHVALPMRPSQRTLOKQG 503

# RESULT 71

MRB  
 Ig mu chain C region, secreted form - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 13-Aug-1986 #sequence\_revision 13-Aug-1986 #text\_change 23-Aug-1997  
 C:Accession: A02164  
 R:Bernstein, K.E.; Alexander, C.B.; Reddy, E.P.; Mage, R.G.  
 J:Immunol. 132, 490-495, 1984  
 A:Title: Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of V-Haz a  
 A:Reference number: A02164; MUID:84088930; PMID:6418603  
 A:Accession: A2 A10Ctype  
 A:Residues: 1-458 <BER>  
 A:Molecule type: mRNA  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:21-92/Domain: immunoglobulin homology <IMM1>  
 F:130-203/Domain: immunoglobulin homology <IMM2>  
 F:242-310/Domain: immunoglobulin homology <IMM3>  
 F:349-420/Domain: immunoglobulin homology <IMM4>  
 F:14/Disulfide bonds: interchain (to light chain) #status predicted  
 F:28-90,137-200,249-308,356-418/Disulfide bonds: #status predicted  
 F:46,114,212,261,277,284,445/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:219,457/Disulfide bonds: interchain (to heavy chain) #status predicted  
 F:296/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 13.3%; Score 360; DB 1; Length 458;  
 Best Local Similarity 23.8%; Pred. No. 3.5e-14;  
 Matches 118; Conservative 76; Mismatches 145; Indels 156; Gaps 22;

QY 21 AATGCKVVLGKKGDVVELTCTASQ--KSIQFHW--KSNQIKILGNGSLTKGPKSL 76  
 DB 17 ALTDGVLVWG-----CLARDPLPSSTVFSFKNKNSI----- 50  
 QY 77 NDRAUSRSRLMDQGNFPLIIKILK-----IEDDTY-ICEVEDQKEVQL 121

DB 51 -----SRTV---RTFVVKRGDKYMATSOVLVPSKDLQGTBEYLCKVQHSNRRDLR 102  
 QY 122 VEGLTANSQDTHLQGSILTLTLESPGSSPSVQGR-----SPR-- 159  
 DB 103 V-----SEFVSELPNNVSVFIPPDSSFGSGSTRKSRLLCOATGFSFKOI 147

QY 160 -----GKNIQGG-----KTLVSQLELDGSGTW-----TCTV-----L 187  
 DB 148 SVSMRLDQKVBSSGLTKPVEAETGAGAPATFISISMLTTESDVLQSLYTCRDHNGI 207  
 QY 188 QNOKVEFKIDIVCPAPRPPKSCDKHTCPBELLGSPVFLPPPKDTLMISRTPEVTCV 247  
 DB 208 FFDKKNVSSSECSSTPSP-----GIQFPIAPSPADT-FLSKSARLICT 250  
 QY 248 VVDVSHEDPEVKENMYVDGVEVHNAKTKPR--EQYNSTYRVSULTVLHODMLNCK 302  
 DB 251 VYDLITTYG-SLNISSV-----ASHNKGALDTHNNTESHPPNATFSMAGASVCAEDWESGE 304  
 QY 303 EYKCVSNKALPAPIEKTISKAKGQPRE--POVYTLPPSRDELT-KNQVSLTCLVYGFYPS 360  
 DB 305 QPTCTVTHADLPFLPKHTISKREVAKIPRAVYULPPAREQLVRESATVTCVYGFSPA 364  
 QY 361 DIAVESNNGQP--ENNYKTPPVLD--SDGSFFLYSKLYDKSRMOQGNFVSCSVMEHA 416  
 DB 365 DVFVQMOQGRQLSSDKVTTSAPADEPQAPGLYFTHSTLVTEDMNGSEFTFCVVGHBA 424  
 QY 417 LHNHYTKSLSPG 431  
 DB 425 LPHMYTERTVDKSTG 439

# RESULT 72

160082  
 CD4 receptor - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 23-Jul-1999  
 C:Accession: I60082  
 R:Zverev, V.V.; Sidorov, A.V.; Nedospasov, S.A.; Malinbova, V.V.; Udalova, I.A.; Andzha  
 Vopr. Virolog. 40, 100-102, 1995  
 A:Title: [Nucleotide sequence of two exons of the human T-lymphocyte CD4 receptor gene].  
 A:Reference number: I60082; MUID:95407135; PMID:7676667  
 A:Accession: I60082  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-71 <RES>  
 A:Cross-references: GB:S79267; NID:G1086922; PIDN:AAAB35273.1; PID:G1086923  
 C:Genetics:  
 A:Introns: 17/1  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

Query Match 13.2%; Score 357; DB 2; Length 71;  
 Best Local Similarity 98.6%; Pred. No. 5.7e-15;  
 Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPEFHLILVQLALPPAATGKNVILGKKGDVVELTCTASQKSIQFHWKNSNIOK 60  
 DB 1 MNRGVPEFHLILVQLALPPAATGKNVILGKKGDVVELTCTASQKSIQFHWKNSNIOK 60  
 QY 61 ILGNGSEFLTK 71  
 DB 61 ILGNGSEFLTK 71

# RESULT 73

S03961  
 Ig mu chain C region - house shrew (fragment)  
 C:Species: Suncus murinus (house shrew)  
 C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 21-Jan-2000  
 C:Accession: S03961  
 R:Ishiguro, H.; Ichihara, Y.; Namikawa, T.; Nagatsu, T.; Kurosawa, Y.  
 FBS Lett. 247, 317-322, 1989  
 A:Title: Nucleotide sequence of Suncus murinus immunoglobulin mu gene and comparison wit  
 A:Reference number: S03961; MUID:89232144; PMID:2497033

A:Accession: S03961  
 A:Molecule type: DNA  
 A:Residues: 1-457 <ISH>  
 A:Cross-references: EMBL:X13920  
 C:Genetics:  
 A:Introns: 106/1; 221/1; 327/1  
 C:Superfamily: Immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:241-309/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 354.5; DB 2; Length 457;  
 Best Local Similarity 25.5%; Pred. No. 7.4e-14;  
 Matches 122; Conservative 71; Mismatches 163; Indels 123; Gaps 19;

OY 19 LPAAFGKGVKVLGKGGDTVELTCTASQ--KKSIOFHMNSNOIKTLGNQGSFLTKGPEKL 76  
 DB 17 LPDEIQ-----VTLGCLARDFLPRVITSMKFKNSSI----- 49  
 OY 77 NDRADSRSLMDQGNFLIKNLK-----IEDSDTYI-CEVEDOKEVOL 121  
 DB 50 -----SSQNIY--NFPEVFTGKYMATSQVLLPSTALIGSTDDYITHTHTTGKEXKK 101  
 OY 122 VFGLTANSDDLHLCQSLTLTLSPSSPSVQCRSPGKNIQGGKTLVSQLELDSC- 180  
 DB 102 V-----ELQVTPLELPNVISIFV--PPR--NSFSGNHPRTSLICQASGF 141  
 OY 181 ---TWCTVLONOKKVEFKIDIVPCAPPEKSCDKT-----HTCP 217  
 DB 142 SPRTVMSWLRGGEVQSLVSTSAVEAPKSGCTTTRVIRSLTTENEMLSOREFTCQ 201  
 OY 218 EL-----LG-----GPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEV 258  
 DB 202 ALHKGLTFQKNVSVCMGDDTSTGISVFLPPTFAN-IFLQSGQLTLVLGATYD-SL 259  
 OY 259 KFNMYDVGVNNAATKTRREOCYNSTYVVSVLTVLHODMLNGEKYKCKSNKALPAIE 318  
 DB 260 DISMSRONGEALQTHVNISEHPNSTFPAKHASVCREWESGKEFTCTVQHSDLPSPLK 319  
 OY 319 KTISAKAGPRE-POVYTLPPSRDEL-TKNQVSLCLVKGYPSPDINEMESNGOP--EN 374  
 DB 320 QSLSPKQVANDPSPVFLPQEQOLKRESASTLCVXDSPSPVDFQWQHGOQPDVDPK 379  
 OY 375 NYKTTTPVLD--SDGSFPLSKLTVDKSRMOQGNVFCSVVMEALNNHYTOKSLSLSPG 431  
 DB 380 HVTYNSNPPRPQNPGLYVHVSHILTVSEKDMSSGSFSCVVGHEALPLSTVEKAVDKNSG 438

## RESULT 74

S25644  
 Ig mu chain C region - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 20-Feb-1995 #sequence\_revision 30-Jan-1998 #text\_change 21-Jan-2000  
 C:Accession: S25644  
 R:Parter, K.; Bugeon, L.; Soullion, J.P.  
 Submitted to the EMBL Data Library, September 1992  
 A:Reference number: S25644  
 A:Accession: S25644  
 A:Molecule type: mRNA  
 A:Residues: 1-343 <PAR>  
 A:Cross-references: EMBL:X68312; NID:956461; PIDN:CAA48392.1; PID:9818025  
 A:Experimental source: spleen  
 C:Genetics:  
 A:Map position: 6  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:234-305/Domain: immunoglobulin homology <IMM>

Query Match 13.1%; Score 353; DB 2; Length 343;  
 Best Local Similarity 30.0%; Pred. No. 6.4e-14;  
 Matches 87; Conservative 66; Mismatches 107; Indels 30; Gaps 10;  
 OY 152 SVQGRSPGKNIQGGKTLVSQLELDSCGTTCTVLONOKKVEFKIDIVPCAPPEKSCD 211

DB 55 TVEAKGSRPQTYKIVISTLTITESDMLNANFTCRV--DHRGLTFWKVNSSTCAASPST-- 110  
 OY 212 KTHCPBLGSPSFLPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYDGVNHN 271  
 DB 111 -----DILAFP---IPSPAD-IFLTKSAKSLCTVNTLATYD-TLNISWSSKSGEPL 158  
 OY 272 AKTPREQYNSTYVVSVLTVLHODMLNGEKYCKSNKALPAPIEKTISKAGPRE- 330  
 DB 159 TMTIMESHNGFSAAGVAVSCMEDNDNRKEFCVTHHDLPSPOKFFISK-----FNEV 214  
 OY 331 ---POVYTLPPSRDEL-TKNQVSLTCLVKGFPSPDIAVEMESNGOP--ENNYKTPPVL 383  
 DB 215 AKHPAPAYLLPPAREQLITRESATVTCVKGFPSPDIFVQMLRGQPLSDPKYVTSAPMP 274  
 OY 384 D--SDGSFPLSKLTVDKSRMOQGNVFCSVVMEALNNHYTOKSLSLSPG 431  
 DB 275 EPGAGLYFTHSILTVEEEMNGEYTCVVGHEALPHMYERTVDKSTG 324

## RESULT 75

IMHMY  
 Ig mu chain C region - golden hamster  
 C:Species: Mesocricetus auratus (golden hamster)  
 C>Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Aug-1996  
 C:Accession: A02168  
 R:McGuire, K.L.; Duncan, W.R.; Tucker, P.W.  
 Nucleic Acids Res. 13, 5611-5628, 1985  
 A:Title: Phylogenetic conservation of immunoglobulin heavy chains: direct comparison of  
 A:Reference number: A02168; MUID:85297761; PMID:2994005  
 A:Accession: A02168  
 A:Molecule type: DNA  
 A:Residues: 1-454 <MCG>  
 A:Note: the sequence was determined from the germline gene  
 C:Genetics:  
 A:Introns: 105/3; 218/3; 324/3  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-90/Domain: immunoglobulin homology <IMM1>  
 F:128-200/Domain: immunoglobulin homology <IMM2>  
 F:238-306/Domain: immunoglobulin homology <IMM3>  
 F:345-416/Domain: immunoglobulin homology <IMM4>  
 F:45-112, 139, 192, 210, 238, 257, 280, 326, 441/Binding site: carbohydrate (Aan) (covalent) #str:  
 F:135-198, 245-304, 352-414/Disulfide bonds: #status predicted

Query Match 12.9%; Score 349; DB 1; Length 454;  
 Best Local Similarity 24.6%; Pred. No. 1.6e-13;  
 Matches 113; Conservative 82; Mismatches 154; Indels 110; Gaps 19;

OY 37 VELTCTASQ--KKSIOFHMNSNOIKTLGNQ--SFLTKPSKLNDRADSRSLMDQGNF 92  
 DB 23 VAMGCLARDFLPSSISFSNMYQNKSEV--NQGVRTPEPTL-----FMGKVAATSGVFL 73  
 OY 93 PLIIKULKIEDSPY-ICEVE--DQKEVOLLVFGLTANSDDLHLCQSLTLTLSPGS 149  
 DB 74 P---PKSVLESGDEYLVCVKNHGNTKDLKPLPGVT-----ENNPNV 113  
 OY 150 SPVSVOCR-----SPR-----GKNIOGG----- 166  
 DB 114 SVFVPSRDASGPAPRKSRLFCEASNPSPKQITVSWLRDQKPYKSGFTTTPVPEDRGSG 173  
 OY 167 -----KTLVSQLELDSCGTTCTVLONOKKVEFKIDIVPCAPPEKSCDKTHTTCPEL 219  
 DB 174 PRITYKIVISTLTITESDMLNANFTCRV--DHRGLTFWKVNSSTCAASPST----- 221  
 OY 220 LGPSVSLPPPKKDT--LMSRTPEVTCVVVDVSHEDPEVKFNMYDGVNNAATKPRE 278  
 DB 222 ----DIOAFPIPSVGFIFLNSKATVLCVNTLATYD-TLNISWSSKSGEPLFTKYLTE 276  
 OY 279 EQYNSTYVVSVLTVLHODMLNGEKYCKSNKALPAPIEKTISKAGPRE-EPQVYTL 337

Db 277 SHPNGTFAIGBANVCVEDMDSGKEFVCTVTRHDLPSPOKFEISKPRGMNKTTPAVYQOP 336  
 Qy 338 PSRDEL-TKNQVSLTCLVGVFPSPDIAYEWESNGOP--ENNKTTPPVLDSDGS--FFLY 392  
 Db 337 LAREQILLESATVCLVGVSPADIVMORGLSDQKVTSAFMKEPQAPHLVFTH 396  
 Qy 393 SKLTVDKSRMGOQGVNFCVMEALHNHYTKSLSPG 431  
 Db 397 SVLTJTEEBMNGEFTYTCVGHGHALPHMVTERTVDRSTG 435

## RESULT 76

C31933

Ig mu chain C region - African clawed frog (fragment)

C/Species: Xenopus laevis (African clawed frog)

C/Date: 31-Mar-1990 #sequence\_revision 30-Jun-1992 #text\_change 20-Sep-1999

C/Accession: C31933

R/Schwager, J.; Mikoyak, C.A.; Steiner, L.A.

Proc. Natl. Acad. Sci. U.S.A. 85, 2245-2249, 1988

A/Title: Amino acid sequence of heavy chain from Xenopus laevis IgM deduced from cDNA se

A/Reference number: A94192; MUID:88176921; PMID:2451244

A/Accession: C31933

A/Molecule type: mRNA

A/Residues: 1-453 &lt;SCH&gt;

C/Superfamily: Immunoglobulin C region; immunoglobulin homology

C/Keywords: immunoglobulin

Query Match 12.8%; Score 345.5; DB 2; Length 453;  
 Best Local Similarity 25.8%; Pred. No. 2.5e-13;  
 Matches 119; Conservative 77; Mismatches 189; Indels 77; Gaps 18;

Qy 31 GKSGDTVELTCTASQ--KKSIOFMKNSQIKILNQGSF---LTGKPSKILNDRADSR 84  
 Db 15 GSGMPVITIGCLAKDFLPETISFTWGDKNNASYSTGLKYKPVMOSSGYTASQVNVAS 74  
 Qy 85 SLMDQGNFLIINKLIEEDSTYICEDVQKEVQ-----LTV 122  
 Db 75 AYWD--NIQFPCNAHLDLT---IKVELKDPVKVEKPVVSIHPPSKDALALNESLFI 129  
 Qy 123 FGLTAN-SDTHLLQGSGLTLTLESPPGSSPVQCRSP--RGKNIQGGKTLISVQLELD 178  
 Db 130 VLATFTPTTHI-----VIKMLKNGQTTGEGVREVPEDKKRGVATSYLSITREKMDL 184  
 Qy 179 SGTWCTVYQNGKVEFKIDIVPCAPBEKSCDKTTCGELGGR--VFLPPPKQTL 236  
 Db 185 DTLVSCVV-----EHAESGSLQEKMSKSLMCTPTPTTSIQVITTPS-LESI 232  
 Qy 237 MISRPPEVTCVVVDVSHEDPEVKFNNY--VDGVEVNHAKTKPREEOYNS--TYRVVSVLTV 293  
 Db 233 PEKSKTTLTCLVSNANSEDLSISWFKKSGTOEITLKTGLDALYNDKRTYSVKGCTTV 292  
 Qy 294 LHDWLNKGYKCKVSNKALPAPIETKISKANGQPREPOVYTLPSRDELTKNQ-VSLTC 352  
 Db 293 CADEMNNDK-FVCKVETHELASWKEVFLPEKEGEVYTPVYVFPPLBELSKREFTALTC 351  
 Qy 353 LYKGFPSPIAYEM--ESNGQPENNYKTP-----PVLSDGSPFLYKLVVDKSRMQ 404  
 Db 352 LYKGPSPEI FVKMLHKNVAVPKQNYINTSINDELLPKQKSGFLYSLHTIDIKMDWA 411  
 Qy 405 GNVFSCVMEALHNHYTKSLSPG-----LQLETC 438  
 Db 412 GDSFSCVGHESLPQLTORSIDKSSGKFTNNVNSVLSDTC 453

## RESULT 77

B30503

Ig gamma-2a chain C region (55.7A12) - mouse (fragment)

C/Species: Mus musculus (mouse mouse)

C/Date: 31-Mar-1989 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999

C/Accession: B30503

R/Gilmore, G.L.; Bard, J.A.; Birnstein, B.K.

J. Immunol. 141, 1754-1761, 1988

A/Title: DNA rearrangements affecting both variable and constant regions of Ig H chain g

A/Reference number: A30503; MUID:88315788; PMID:2842402

A/Accession: B30503

A/Molecule type: mRNA

A/Residues: 1-112 &lt;GIL&gt;

A/Cross-References: GB:M21925

A/Experimental source: myeloma cell line MPC11

A/Note: the authors translated the codon GAG for residue 11 as Ser

C/Genetics:

A/Introns: 100/3

C/Superfamily: immunoglobulin C region; immunoglobulin homology

C/Keywords: immunoglobulin

F:1-70/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 12.3%; Score 311.5; DB 2; Length 112;  
 Best Local Similarity 50.0%; Pred. No. 3.2e-12;  
 Matches 66; Conservative 17; Mismatches 22; Indels 27; Gaps 2;

Qy 239 SRTEPVCVVVDVSHEDPEVKFNNYVDGVEVNHAKTKPREEOYNSYTVSVTLVLDHW 298  
 Db 1 SLTPKVTCTVVDVSDDDVDQISWPNVVEHTAQOTHRREDYNTIRVSTLPQHQM 60  
 Qy 299 LNKGYKCKVSNKALPAPIETKISKAKQPREPOVYTLPSRDELTKNQVSLTCLVKGY 358  
 Db 61 MSGKEFKCKVNKKDLPAPIERTISKPE-----SCSLTAWGLGW- 100  
 Qy 359 PSDIAYEWESNG 370  
 Db 101 -----WTSNG 105

## RESULT 78

I50731

Ig heavy chain - nurse shark

C/Species: Ginglymostoma cirratum (nurse shark)

C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C/Accession: I50731

R/Vazquez, M.; Mizuki, N.; Flajnik, M.F.; McKinney, E.C.; Kasahara, M.

A/Title: Nucleotide sequence of a nurse shark immunoglobulin heavy chain cDNA clone.

A/Reference number: I50731; MUID:92357056; PMID:1495502

A/Accession: I50731

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-577 &lt;VAA&gt;

A/Cross-References: GB:M2851; NID:9213264; PIDN:AAA50817.1; PID:9213265

C/Genetics:

A/Gene: IGH

C/Superfamily: immunoglobulin C region; immunoglobulin homology

F:469-539/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 12.0%; Score 325.5; DB 2; Length 577;  
 Best Local Similarity 24.3%; Pred. No. 5.1e-12;  
 Matches 142; Conservative 83; Mismatches 165; Indels 195; Gaps 28;

Qy 10 LLLVQLALPRAATGKNV-----LGKKGDVELTCTASQKSIQFMKNSQ--IKIL 62  
 Db 6 IFLSLALLALPCVQSEITLIDPEATGHPGSLTLC---KINGFNVGSSSMYMKOV 60  
 Qy 63 GNOG-----SFLTGKPSKLANDRSDRSIMDOGN--FPLIKNLKIEDSDTYICEV 111  
 Db 61 PGQGLEWIVYVYSSSMNYPAPAIKDRFTRAK---DTSNNIFLMBRSVKIDTALYYC-- 115  
 Qy 112 EDQKEEVOLLVGLTANSPTHLQGSILTLTLESPPGSSPV-----QGRS----- 157  
 Db 116 -----TRMSGYEYLGGHSGYWGQGTWVTVTATP--SSPTLYGLVSSCOQGNIDSVLY 168  
 Qy 158 -----PRGKNIQGGKTLISVQL----- 174  
 Db 169 GCLAMDYPDVAVSWTKKHGOLLITTVOTYPSVRNKKGYTUS-SQLALLIESDAECDOIS 227  
 Qy 175 -ELQDSG-----TWCT 185  
 Db 228 CEVRHSGSDKSTGMCPCPDGFTALLTVSSSEIESRKAIIYCVSISDFHSKISVTV--- 284



```

QY 186 VLQONKAYEFKIDIVP-CPA-----PEPKSCDK-THTC-----PEL-----LGG 222
DB 285 -LKNGRSVDSGIFSTPVECANNGNFSVTSLRLRPVAMWEDRAVYTCQVAYKEVIOQSWNTG 343
QY 223 PSVF-----LFPKPKDITMISRTPEVTCVVVDVSHEDPEVKENNYVDGVEVHNAKT 274
DB 344 PVSSECHGYTAKILPP-PVEQVLTLEATVTLTCV---VSNLSGVNFTWLOD-----EKT 393
QY 275 KPREQVNS---TYRVSVTLVLHODMNLGKGYKCKVKNKALPAIEKTIKAK-GOPRE 330
DB 334 LKSELAHNSGHSDAISKLDISTAMLSSEVFECVNHQYLPPLRDSIHKERLENLE 453
QY 331 POVYTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGOPEN--NKTTPPYLSDG 387
DB 454 PVSVTLPTTEELSNQRFSLTCLVLRGRFPREIFPKMTTNNKPVNPSVYKNTVTAESDN 513
QY 368 -SFFLYSKLTVDKSNWQGNVFSCVMEALHNITQKSLSPG 431
DB 514 TSFFLYSLSLIAERWASGASYSYCVGHEAIPLKTIINRTVKKSSG 558

```

## RESULT 79

```

S31436
I9 upslon chain - axolotl (fragment)
C:Species: Ambystoma mexicanum (axolotl)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999
C:Accession: S31436
R:Feilich, J.S.; Miles, M.V.; Schwager, J.; Charlemagne, J.
Submitted to the EMBL Data Library, November 1992
A:Description: cDNA sequence of Ambystoma mexicanum upslon heavy I9y chain.
A:Reference number: S31436
A:Accession: S31436
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-433 <FEU>
A:Cross-References: EMBL:X69492; NID:G62420; PID:G62421
C:Superfamily: Immunoglobulin C region, immunoglobulin homology

```

```

Query Match 11.9%; Score 321.5; DB 2; Length 433;
Best Local Similarity 24.2%; Pred. No. 6.3e-12;
Matches 111; Conservative 65; Mismatches 146; Indels 137; Gaps 20;

QY 28 VLGKKGDTVEL-----TCTASQKKSIOFMKNKNQIKILGNGSFLTKGPS 74
DB 28 LVTGLGPRVEVSWKAGSTMPGKTPFAQQQAAL--STSSQRIAYE----- 76
QY 75 KLNDRAISRRLMDQGNFPLIILKLIKEDSDTYI-----CEVEDQKEEYQL----- 120
DB 77 -----WGTSYSCSVKH--KPTSTEIHKITTSABCKKATKSPVQVQLQSSCA 121
QY 121 -----LVFGLT-----ANSDTHLQGSLLTLESPPGSSPSVQCRRSPR 159
DB 122 DTDGNGSIELVCLISGVTPDNIQVRLVNDKAPRIQGGT-----SPR 163
QY 160 GKNIGG---GKTLVSQLELDSDGTTCTVLQONKAYEFKIDIVPCAPAPKPSCKDTHT 215
DB 164 QKDGGTSTTSQIINVTSDMASGDKYTC-----KVEH-----PATSSRAEDITHN 209
QY 216 CPELLIG--GPSVFLFPPPKDITMISRTPEVTCVVVDVSHEDPEVKENNYVDGVEVHNAK 273
DB 210 CADSGTPYQKVFLLAPKARD-LYIANQPVVICTIKKENSD-SLSVTM-----KRR 259
QY 274 TKPRE-----EGY---NSTRVSVTLVLHODMNLGKGYKCKVKNKALPAIEKTIKAK 325
DB 260 EGPEEAAVISEQYIDSDGFTAMSLNTIKMEWEGDFTCKVKHFDLPFLPSVSKSPPT 319
QY 326 GQPREPOVYTLPPSRDELTK-NQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTPPV 382
DB 320 GRSFAPRTVYVAPHEMLANDVSLTCLVKSFSFDDIYIQMKQCKSVIPSDKYVSMRPR 379
QY 383 LDSD---GSFFLYSKLTVDKSNWQGNVFSCVMEAL 417

```

```

DB 380 QEAGTAGLGTFFSYSMULTIQKSDMDKRETFVAAHSAV 418

```

## RESULT 80

```

A39016
T-cell surface glycoprotein CD7 precursor - human
N:Alternate names: T-cell leukemia antigen
C:Species: Homo sapiens (man)
C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 05-Nov-1999
C:Accession: A39016; S03520
R:Schenberg, L.E.; Fleener, D.E.; Kurtzberg, J.; Haynes, B.F.; Kaufman, R.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 603-607, 1991
A:Title: Isolation and characterization of the genomic human CD7 gene: structural simila
A:Reference number: A39016; MUID:91110576; PMID:11703303
A:Accession: A39016
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-240 <SCH>
A:Cross-References: GB:M37271; NID:G180163; PIDN:AAA51953.1; PID:G180164
R:Aruffo, A.; Seed, B.
EMBO J. 6, 3313-3316, 1987
A:Title: Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a COS cell expr
A:Reference number: S03520; MUID:88111517; PMID:3501369
A:Accession: S03520
A:Molecule type: mRNA
A:Residues: 1-240 <ARU>
A:Cross-References: EMBL:X06180; NID:G29819; PIDN:CAA29546.1; PID:G732757
C:Genetics:
A:Gene: GDB:CD7
A:Cross-References: GDB:119770; OMIM:186820
A:Map position: 17q25.2-17q25.3
A:Insertions: 28/1
C:Keywords: glycoprotein; membrane protein; surface antigen; tandem repeat
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-240/Product: T-cell surface glycoprotein CD7 #status predicted <MAT>
F:145-180/Region: 9-residue repeats (X-P-P-X-A-S-A-L-P)

```

```

Query Match 11.7%; Score 316.5; DB 2; Length 240;
Best Local Similarity 41.7%; Pred. No. 6.2e-12;
Matches 90; Conservative 22; Mismatches 54; Indels 49; Gaps 11;

```

```

QY 321 ISKAKGP-----REPQVYTLPPSRDELTKNOVSLTCL---VKGFY-----PS 360
DB 14 LALARGPLGALAAQVQSGPHCTTVPGA-----SVNICTSTSGGRGIVLRQLGPQ 67
QY 361 DIAVWESNGQPENNYTPTPVLDSGFFLYSKLTVDKSRMNGN--VSCSVMEALH 418
DB 68 DI-LYEDGVVPTTDRFRGRIT-DPSGS---QDNLTITMRLLQSDGTTCQATIEV-- 120
QY 419 NHYTQKSLSPGLQLD--ETCAEADQGEIDGLMTTPPRAALPAPPTGSALPDPTAS 476
DB 121 NVYGSGLVLTVEGQSGMHCSDA-----PPRAASLPAPPTGSALPDPTAS 168
QY 477 ALPPRAASLPALAVISFLGLIGLGVACVLART 512
DB 169 ALPPRAASLPALAVISFLGLIGLGVACVLARTQ 204

```

## RESULT 81

```

I9 mu chain C region (clone 12022) - horn shark (fragment)
C:Species: Heterodontus francisci (horn shark)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: S00980
R:Kokubo, F.; Hinds, K.; Litman, R.; Shambloct, M.J.; Litman, G.W.
EMBO J. 7, 1979-1988, 1988
A:Title: Complete structure and organization of immunoglobulin heavy chain constant regic
A:Reference number: S00980; MUID:88328985; PMID:3138109
A:Accession: S00980
A:Molecule type: mRNA
A:Residues: 1-438 <KOK>
A:Cross-References: EMBL:X07784; NID:G63963; PIDN:CAA30617.1; PID:G63964
A>Note: the sequence was determined from the differentiated gene

```

C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin

F1-438/Domain: C region <CRE>

F120-87/Domain: immunoglobulin homology <IM1>

F123-190/Domain: immunoglobulin homology <IM2>

F128-291/Domain: immunoglobulin homology <IM3>

F130-400/Domain: immunoglobulin homology <IM4>

F166,200,245,275,374,411,415,423/Binding site: carbohydrate (asn) (covalent) #status pr

Query Match 11.5%; Score 312; DB 1; Length 438;

Best Local Similarity 25.4%; Pred. No. 2.3e-11;

Matches 113; Conservative 67; Mismatches 155; Indels 110; Gaps 17;

```

QY 23 TGGNKV---VLGGKDTVELTCTASOKSIQFHMKNSNQIKILGNOSGLTGPSTLND 79
DB 49 TTGLKIYPSVLNKKGT-----YTRSSQLTITSEB-----VGSXKIYCE 86
QY 80 ADSRSLMDQGNPPLIKNLKIEDSDTYICEVEDQKEEYQLLVFGLTANSQTHL----- 133
DB 87 VARGESLW-----IKETLDCKGDIVPTVILITSSSEITSRRA 126
QY 134 -----LQGSLLTL-----LESPPGSSPSVQCRSPRGNIOGKTLVSQLELQD 178
DB 127 TVLCSTIDFHPESITVSWLKDGQPMDSGFVTSPTEVNG-----NFSATRLTVPAGEWS 182
QY 179 SGTWTCTVQLQNKCKEFK--DIVCPAPBPSCDKTHTCPBLGGPSVFLPPPKQTL 236
DB 183 NTVTTCQVHAQETOSRNITGSOVPCS-----IGDPVTKLPPSTIEQVL 226
QY 237 MSRPPEVTVVVDVSHEDPEVKFMYVDGVEVHNAKTPREEQY-----NSTRYVSVL 291
DB 227 -LEATVTLTLCV---VSNADYGVNWS-----TOEKPLSEIAVQGEDSDVISTV 274
QY 292 TVLHODMLNGEKYCKVSKNALPAPIEKTISKAKGP-REPOV-YTLPPSRDELTKNQVS 349
DB 275 NSTQWMLSGAEFYCVVSHQDLPTPLASIHKEEVLDAREPVSVALPPAEVDVSNQRF 334
QY 350 LTCLVKGFPYSDIAWESNNGPEN--NYKTPPVLDSPG-SFFLYSKLTVDSKMOQN 406
DB 335 LTCLVKGFPYSDIAWESNNGPEN--NYKTPPVLDSPG-SFFLYSKLTVDSKMOQN 394
QY 407 VFGSCVMEHALNHNYSKLSLSPG 431
DB 395 SYSCVVGHEAIPDKIINRTVKNSSG 419

```

#### RESULT 82

Ig mu chain - Leptosteus oboeus (fragment)

C:Species: Leptosteus oboeus

C>Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 03-Nov-2000

C:Accession: 150830

Dev. Comp. Immunol. 19, 133-144, 1995

A>Title: cDNA sequences and organization of IgM heavy chain genes in two holostean fish.

A:Reference number: 150006; PMID:96039719; PMID:7556802

A:Accession: 150830

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-474 <HL>

A:Cross-references: EMBL:U12455; NID:9529949; PIDN:AACS9688.1; PID:9529950

A:Gene: IGM

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 11.4%; Score 308.5; DB 2; Length 474;

Best Local Similarity 24.2%; Pred. No. 4.2e-11;

Matches 116; Conservative 62; Mismatches 171; Indels 131; Gaps 20;

31 GKKGDTVELTCTAS--QKSIQFHMKNSNQIKILGNOSGLTGPSTLNDRAISRSLMD 88

DB 50 GTSGDVVALGLATGFLPDLSTLFKMTDSTDKEL-----TPPRKPSVLNGETYSSTS--- 101

QY 89 QGNFPLIKNLKIEDSD-----TYICEVEDQKEEYQLLVFGLTANSQTHLQGSLLTL 143

DB 102 -----QSLPSPDMNSGKAFCEAKHPQGDVYL-----HL----- 132

QY 144 ESPPGSSPSVQCRSPRGNIOGKTLVSQLELQDGTWCTVLO-NOKKVEFK----- 196

DB 133 -TPPVPAPSVLNLNP-----SLEEPQNTATLVCVRSGFSPKTHFFKMRGNT 180

QY 197 -----IDVPCPAPPEK-----SCDKTH-----TC 216

DB 181 KLDEGVNTIPATVDDKTKYSASLLTVEKQWKSAPACEVHTGTVLKNITYSRRC 240

QY 217 PELGGSPVFLPPPKQKTMISRTPEVTC---VVVDVSHEDPEVKFMYVDGVEVHNAK 273

DB 241 QETV---KVLEPPTEBQ-EVKKTATLTCRIALVSTS---DVSMTSSGGKFL--AA 290

QY 274 TKPREQVNSTYRVSVLTVLHODMLNGEKYCKVSN-KALPAPIEKTISKAKGPREFQ 332

DB 291 GAPERGHGCGKVAASRVSVLEKRTGTEYKCIYSHDSFPTLTKTYKQDIATKIPS 350

QY 333 VYTLPPSRDE--LYKNQVSLTCLVKGFPYSDIAEW--ESNGQPNNTKTPPVLDSPGS 388

DB 351 VFLAPSTEENSTDEVTLTFCFVDFFPKDIYISMLADSIYDKHTHTVTDLIPSHDGA 410

QY 369 FFLYKSLTVDSKMOQNVFSCVMEHA-----LHNHYTKSLSLSGQLDLETC 438

DB 411 YSVSKYTISSSDMNSGTMVSCAHHETAPLPVSVITRTDSTGKATLVNFTLPTDC 470

#### RESULT 83

T-cell surface glycoprotein CD4 (allele 1) - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 20-Feb-1995 #sequence\_revision 19-Apr-1996 #text\_change 21-Jan-2000

C:Accession: 147131; S21461

J. Immunol. 151, 1365-1370, 1993

A>Title: Extensive allelic polymorphism in the CD4-like region of the miniature swine

A:Reference number: 147131; PMID:93329116; PMID:8335933

A:Accession: 147131

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-99 <GU2>

A:Cross-references: EMBL:X65629; NID:91928; PIDN:CAA46563.1; PID:9388232

C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C:Keywords: glycoprotein; T-cell

F1-81/Domain: immunoglobulin homology <IMW>

Query Match 11.3%; Score 305.5; DB 2; Length 99;

Best Local Similarity 60.2%; Pred. No. 9.7e-12;

Matches 59; Conservative 16; Mismatches 22; Indels 1; Gaps 1;

32 KKGDTVELTCTASOKSIQFHMKNSNQIKILGNOSGLTGPSTLNDRAISRSLMDPG 90

DB 1 KAGDLAELPCHSOKKNLFFMKNSNQIKILGHSFMTHTASVETLSKLSKMMWHDG 60

QY 91 NFPLIKNLKIEDSDTYICEVEDQKEEYQLLVFGLTAN 128

DB 61 SFPLIKNLLEVDSGIVCEVEDKRIEYQLVFRITAS 98

RESULT 84

Ig mu chain C region, membrane-bound (clone 3050) - horn shark

C:Species: Heterodontus francisci (horn shark)

C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Aug-1996

C:Accession: S01854; G2716; A46530

R:Kokubu, F.; Hinder, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.

EMBO J. 7, 1979-1988, 1988

A>Title: Complete structure and organization of immunoglobulin heavy chain constant reg;

A:Reference number: S00980; PMID:88328985; PMID:3138109

```

A:Accession: S01854
A:Molecule type: DNA
A:Residues: 1-461 <KOK>
A:Cross-references: EMBL:X07781
A>Note: the sequence was determined from the germ-line gene
R:Kokubu, F.; Hinde, K.; Litman, R.; Shambloet, M.J.; Litman, G.W.
Proc. Natl. Acad. Sci. U.S.A. 84, 5868-5872, 1987
A>Title: Extensive families of constant region genes in a phylogenetically primitive vert
A:Reference number: A32716; MUID:87289703; PMID:3475706
A:Accession: C32716
A:Molecule type: DNA
A:Residues: 1-59 <KO2>
A:Cross-references: GB:M17186
A:Genetics:
A:Introns: 100/1; 206/1; 309/1; 419/1; 459/3
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)
chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 1a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobu
F:120-87/Domain: immunoglobulin homology <IMM2>
F:123-190/Domain: immunoglobulin homology <IMM3>
F:228-291/Domain: immunoglobulin homology <IMM4>
F:330-400/Domain: immunoglobulin homology <IMM4>
F:438-458/Domain: transmembrane #status predicted <TM>
F:27-85,130-188,235-289,337-398/Disulfide bonds: #status predicted
F:164,200,245,374,411,435,437/Binding site: carbohydrate (Asn) (covalent) #status pr
Query Match 11.2%; Score 301.5; DB 1; Length 461;
Best Local Similarity 26.7%; Pred. No. 1e-10;
Matches 121; Conservative 67; Mismatches 161; Indels 105; Gaps 23;
OY 53 WKNSQIKILNQGSLFKGPKSLKDRADSRSLMDQGNFPLIKNLKIEDSDTYICEV- 111
DB 41 WKDNPEITTG-----LKTYPSPVLNKKGTQTS-----SQTITSEVSSKITY-CEVR 88
OY 112 -----EDQKEVQLVFG--LTANSDTLLGQSLTTL- 142
DB 89 RGEVWIKELIDCKGDKVHPVITVILTQSSSEITSRFAVTCISIDFR--PESITVSWL 145
OY 143 -----LESPPGSPSPVQCRSPRGKY--IQGKTLVSQLELDQSGTWTCTVLOKQKYE 195
DB 146 KDQGHMEGFTVSPF--C-----GVNGTSATSRILTPAREFTNKVYTCQV--SHQGTQ 197
OY 196 KIDI---VPCPAPEPKSCDKTHTCPELLGSPVFLFPKPKDTLMTSRPEVTCCVV 251
DB 198 SRNITGSGVPC-----SCN-----DPVTKLPPSIEQVL-LEATYTLTCV--V 237
OY 252 SHEPPEYKFNKYVDCVFNHNAKTKPREQYNSTRVSVLTVLHODMLNGEYKCKYSNK 311
DB 238 SNAFYGVNVSW---TQEQKSLKSEIAVQPGEDADSVISTVNIQTQAWLSGAEFCVNVHQ 294
OY 312 ALPAPIETKISKAKGP-REPQV--YTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESN 369
DB 295 DLPTPLRASIKKEVKDLREPSVSLTSLPAEDVSAQRSLTCLVKGFPPEITVKMTVN 354
OY 370 GQPEN--NYKTPPYLDS--GSFPLSKLTVDKSRMOQGNVFGSVMHEALNNHYTOKSL 426
DB 355 DKSVPNGYKNTVEAENDNSSYFIYSLTSLAEEMASGASVCVGHFAL----- 405
OY 427 SLSPGLQDTCARQDQDELGLGTLTPPPRASAL 460
DB 406 ---PLKTIINRTVKNSSDS--DHIMEDNEESAI 435

RESULT 85
A:Accession: A46507
A:Molecule type: chicken
A:Species: Gallus gallus (chicken)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
A:Accession: A46507
R:Mansikka, A.
J. Immunol. 149, 855-861, 1992
A>Title: Chicken Iga H chains. Implications concerning the evolution of H chain genes.

```

```

A:Reference number: A46507; MUID:92340889; PMID:1634774
A:Accession: A46507
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-585 <MAN>
A:Cross-references: GB:S40610; NID:g251907; PID:g251908
A>Note: Sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:109906, NCBI:P:109907)
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:34-117/Domain: immunoglobulin homology <IMM>
Query Match 11.2%; Score 301.5; DB 2; Length 585;
Best Local Similarity 26.8%; Pred. No. 1.4e-10;
Matches 114; Conservative 60; Mismatches 184; Indels 67; Gaps 22;
OY 19 LPATQGNKVLGKGGDTVELTCTASQKKSIOFMKNSNQIKILNQGSLFKPSKLN 78
DB 197 LPVATGTGYSL-----TVALTVPREQ-----LOGNE--FYCRAQHAATG 234
OY 79 RADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL--LVFGL--TANSDTLLQ 135
DB 235 -ADVKETIGDGVCPITTSKVTLLSDPTQ---EDFERVILVGLVGLPSAGAAIQWLQ 289
OY 136 -GQSLTTLTSSPPGSSPVQCRSPRGKNIGKTLVSQLELDQSGTWTCTVLOKQKVE 194
DB 290 DNEBMT-----PAVESDESQSDCTESGVTQWSKRVNTRKSMEGGAQFGCVTHGALK-- 342
OY 195 FKIDIVCPAPEPKSCDKTHTCPELLGSPVFLFPKPKDTLMTSRPEVTCCVVDSHE 254
DB 343 -----EVTYITVSTDCAT--POL---QVSLIPFLLE-LTVSHNAVTVCVSSNAAA 389
OY 255 DPEYKFNMYVD--GVEFNHNAKTKPREQYNSTRVSVLTVLHODMLNGEYKCKYSNK 311
DB 390 D-GVSVSWSSSSGGGLV-----SQTEDRQADGRVTASPLRVCAENNGGTFPGCSREE 444
OY 312 ALPAPIETKISKAKGP-REPQYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESN 369
DB 445 GV-VVAESIRKETDTPLHAPSVVFPPEAEISLQETATLTCAASSFLPSILLTWTQ 503
OY 370 GQP--ENNYTTPPYLDSQSPF-LYSKLTVDKSRMOQGNVFGSVMHEALNNHYTOKSL 426
DB 504 NQPLSPQNYLIFGB--EKDDGFYSLYSKLKSVEDMQRGDFGCVVGHGDIPLNFIHKS 561
OY 427 SLSPG 431
DB 562 DKNAG 566

RESULT 86
A:Accession: S60266
A:Molecule type: novel antigen receptor precursor - nurse shark
A:Species: Ginglymostoma cirratum (nurse shark)
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 21-Jul-2000
A:Accession: S60266
R:Greenberg, A.S.; Avila, D.; Hughes, M.; McKinney, E.C.; Flajnik, M.F.
Nature 374, 168-173, 1995
A>Title: A new antigen receptor gene family that undergoes rearrangement and extensive s
A:Reference number: S60266; MUID:55183140; PMID:7877689
A:Accession: S60266
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-684 <GRE>
A:Cross-references: EMBL:U18701; NID:g699442; PIDN:AB48195.1; PID:g699443
Query Match 11.2%; Score 301.5; DB 2; Length 684;
Best Local Similarity 25.1%; Pred. No. 1.7e-10;
Matches 107; Conservative 66; Mismatches 161; Indels 92; Gaps 16;
OY 37 VELTCTASQKKSIOFMKNSNQIKILNQG--SFLTKGPKSLMDRADSRSLMDQGNPL 94
DB 321 VETTCSAKQDQSTPVVVKTRKARVEFTKPHLLPSPSEIQTSSATLTCLIRGFDP 380
OY 95 IINKLKIEDSDTYICEVEDQKEVQLVFGLTANSDTLLGQSLTTLTSSPPGSSPVQ 154

```

```

Db      301 KV-----SVSWQKDVSV--VSANTNPETAL--EQDLTFSTRS-----414
Qy      155 CSRPRKNIQGGKTLVSQLELDGSGTWCT-----VQONQKVEFKIDIVPC 202
Db      415 -----LNLTAIVEMKSGAKYCTTASHSPSGSTVRKVRIRNQ-----KVD-----452
Qy      203 PAPEPKSCDKHTCPCELGGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPE-VKRN 261
Db      453 -----CRQTDI-----SVSLKLP-PREEITWQQTATVCEIV---YSDLENIKVF 493
Qy      262 WYVDGEVHNAAKTKPREEQYNSTRYVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTI 321
Db      494 WQVNGVERKKGVETQNPENWSGSKSTIVSKLKVMASEMDSGTEYVGLVEDSELPTPVKASI 553
Qy      322 SKAK-GQPREPOVYTLPPSRDEL-TKNQVSLTCLYKGFPSDIANVESNGO-PENNYKT 378
Db      554 RKAIVSQMHPKCYLHPSTDEIDTENSATLMCLATNFPAPIYVGMAMNDTLDSGYRT 613
Qy      379 TPVLDSDGSPFLYSKLTVDKSRMOQGNVFSQSVMEALH-----NHYTQKSLSPG 431
Db      614 QVDSEKSGSSFTVDRLLTALEMNSDTTYSCLVGHPSLNRDLINSTKNGKPSVAVS 673
Qy      432 LQLDRT 437
Db      674 VVLSDT 679

```

## RESULT 87

```

147175
Ig alpha chain C region - pig (fragment)
C.Species: Sus scrofa domestica (domestic pig)
C.Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C.Accession: I47175
R.Brown, W.R.; Butler, J.E.
Mol. Immunol. 31, 633-642, 1994
A.Title: Characterization of a C alpha gene of swine.
A.Reference number: I47175; MUID:94254897; PMID:7545929
A.Accession: I47175
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-342 <BRO>
A.Cross-references: EMBL:U12594; NID:9555826; PIDN:AAA65943.1; PID:9555827
C.Genetics:
A.Gene: IGACalpha
C.Superfamily: Immunoglobulin C region; immunoglobulin homology
F:232-304/Domain: immunoglobulin homology <IMM>

```

```

Query Match      11.1%; Score 300; DB 2; Length 342;
Best Local Similarity 27.1%; Pred. No. 9e-11;
Matches 96; Conservative 44; Mismatches 138; Indels 76; Gaps 13;

Qy      139 LTLTLESPPG-----SSPSVQCRSPRGKNI-----OGGKTLVSQLEL 176
Db      11 LTLGSEPAQVYVIAVLVDFPSEELVTWSPSREGVIVRFPRAQAGLVTMSQULT 70
Qy      177 QDSGWTCTVLONQKVEFKIDIVCPAPEPKSCDKTH-----TCPE 218
Db      71 -----VEEQCPADQILKCVQHLSKSSQSVNVPCKVLPSPDCPQ 109
Qy      219 LLGGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKRNWYVDGEVHNAAKTKPRE 278
Db      110 CC-KKSLIQPALAD-LTLGSNASTCTLSGLKSE-GVSTWQPSGK-DAVASPR 165
Qy      279 EQYNSTRYVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTIKAKGPREPOVYTLPP 338
Db      166 DSC-GCYSVSITLPGADPMKNGETPSCAASSELKSLATITPKVTVTPRQVHLPP 224
Qy      339 SDELTKNQ-VSLTCLYKGFPSDIANVESNGO-PENNY---KTPPVLDSDGSPFLY 392
Db      225 PSEELALNELVTLTCLVGRFSPKDVLRVWLOGGOELPRKXVYLVWESLPEPGQAITPVAVT 284
Qy      393 SKLTVDKSRMOQGNVFSQSVMEALHNYHTQKSLSPG-----LQLDRTCAEAQ 442

```

```

Db      285 SVLRVDAEDMKGDPFSCVGVHEALPLAFQKTDRLAGKPTHVAVVMAEAE 338

```

## RESULT 88

```

509264
Ig alpha chain C region - rabbit (fragment)
C.Species: Oryctolagus cuniculus (domestic rabbit)
C.Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C.Accession: S09264
R.Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A.Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A.Reference number: S09264; MUID:90076124; PMID:2512120
A.Accession: S09264
A.Status: not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-339 <BUR>
C.Superfamily: Immunoglobulin C region; immunoglobulin homology
C.Keywords: immunoglobulin
F:124-190/Domain: immunoglobulin homology <IMM>

```

```

Query Match      11.0%; Score 297; DB 2; Length 339;
Best Local Similarity 29.0%; Pred. No. 1.3e-10;
Matches 96; Conservative 48; Mismatches 127; Indels 60; Gaps 15;

```

```

Qy      141 LTLSPSSPSVQCRSPRGKNIQ-----GKTLVSQLELDGSGTWCTV 186
Db      10 LSLPCLSGOPVV-----GCLIQFPPLGPLNVKRTISGENVTPPQDLTSGLYTSS 64
Qy      187 LONQKVEF-----KID---IVCPAPEPKSCDKHTCPCEL-LGGPSVFLPPPK 231
Db      65 LNLTDSECPCTVACHVZHEVDRYLPCP-----DTSSSCPCTGSEPSLSLQRPD 117
Qy      232 PKDTLMISRTPEVTCVVDVSHEDP-EVKFNWYVDGEVHNAAKTKPREQYN---STYRV 287
Db      118 LRD-LTLGSDASLTCTLGL--KDPKDAVFTW-----EPTNGNEVPQSPQRDPGCGYSV 169
Qy      288 VSVLTVLHODMLNGEKYCKVSNKALP-APIEKTIKAKGPREPOVYTLPPSRDELTKN 346
Db      170 SSVLPFGCAETVAGTEPCTVTHPEIEGSSLTATIRKDTGSLTPPOVHLPPSEELALN 229
Qy      347 Q-VSLTCLYKGFPSDIANVESNGO-PENNY---KTPPVLDSDGSPFLYSKLTVDKS 400
Db      230 ALVTLTCLVGRFSPKDVLRVWTKGVVVPKQSPFLVWKPFBPGQEPPTVAVTSLRVSAE 289
Qy      401 RMOQGNVFSQSVMEALHNYHTQKSLSPG 431
Db      290 DMNGDSYSVGVGHEGLAHEHTQKTDIDRQAG 320

```

## RESULT 89

```

509269
Ig alpha chain C region - rabbit (fragment)
C.Species: Oryctolagus cuniculus (domestic rabbit)
C.Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C.Accession: S09269
R.Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A.Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A.Reference number: S09269; MUID:90076124; PMID:2512120
A.Accession: S09269
A.Status: not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-357 <BUR>
C.Superfamily: Immunoglobulin C region; immunoglobulin homology
C.Keywords: immunoglobulin
F:142-208/Domain: immunoglobulin homology <IMM>

```

```

Query Match      11.0%; Score 296; DB 2; Length 357;
Best Local Similarity 29.8%; Pred. No. 1.6e-10;
Matches 90; Conservative 40; Mismatches 104; Indels 68; Gaps 14;

```

```

Qy 141 LTLESPSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIV 200
Db 94 VTVPSP-----ECQPTP-----GSPDTTTC----- 115
Qy 201 PCPAPBPKSCDKHTCPRLGSPVLEPPPKKDTLMISRTPEVTCVVDVSHEDPEVKF 260
Db 116 PCPCPSP-SC-----GEPSTLQRPFLRD-LILNSNASLTCTLRGKNEGAV-F 162
Qy 261 NMVYDGEVHNAAKTPREEQYNS-----TYRVSVVTLVHODMLNGKEYCKVSKALP-A 315
Db 163 TW-----EPTNG-NKPVQGSVQSVPCGCTSVSVLPFGAEPMNAGTEFTCTVTHPELEG 216
Qy 316 PIEKTIKAKQCPREPOVYTLTPSRDELTKNQ-VSLTCLVKGYPSPDIAYEWESNG--QP 372
Db 217 PLTKAKISKDTGAILPPQVHLLPPSEELALNELVTLTCLVVGFSFKDVLVYWNKGAVNP 276
Qy 373 ENNV---KTPPVLDSDSFPVLYSKLTVDKSRMOQGNVFCSSVNHIALHNYTKSLSL 429
Db 277 ENSFLVWKPRPEPQEPPTTAVVTSILRVPAEDMWNQESYTCVGHGELAEHFTQRTIDRL 336
Qy 430 PG 431
Db 337 SG 338

```

## RESULT 90

MHCH

Ig mu chain C region - chicken (fragment)

```

C/Species: Gallus gallus (chicken)
C/Date: 03-Aug-1984 #sequence_revision 03-Aug-1984 #text_change 22-Jun-1999
C/Accession: A02170
R/Dahan, A.; Reynaud, C.A.; Weill, J.C.
Nucleic Acids Res. 11, 5381-5389, 1983
A/Title: Nucleotide sequence of the constant region of a chicken mu heavy chain immunogl-
A/Reference number: A02170; MUID:83299221; PMID:6310496
A/Accession: A02170
A/Molecule type: mRNA
A/Residues: 1-367 <DNH>
A/Cross-references: GB:K00389; NID:G212204; PIDN:AAA48923.1; PID:G212205
C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kg)
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a
C/Superfamily: Immunoglobulin C region; immunoglobulin homology
C/Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin
F/49-133/Domain: immunoglobulin homology <IMM2>
F/151-219/Domain: immunoglobulin homology <IMM3>
F/258-329/Domain: immunoglobulin homology <IMM4>
F/347-367/Domain: carboxyl-terminal <CTS>
F/51,119,303,354/Binding site: carboxylate (Asn) (covalent) #status predicted
F/56-111,158-217,265-327/Disulfide bonds: #status predicted
F/126,366/Disulfide bonds: interchain (to heavy chain) #status predicted
F/205/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

```

Query Match 11.0%; Score 296; DB 1; Length 367;

Best Local Similarity 30.1%; Pred. No. 1.7e-10;

Matches 85; Conservative 48; Mismatches 119; Indels 30; Gaps 10;

```

Qy 160 GKNIOGGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIVCPAPBPKSCDKHTC--- 216
Db 87 GPEVVASRISVTSSEMDTGATPSCV-----EGEMRNITSKMEGCGLE 129
Qy 217 PELIGSPVLEPPPKKDTLMISRTPEVTCVVDVSHEDP-EYKFNMYVDGEVHNAAKTK 275
Db 130 PVVQODAIRVITPSPFD-IPISKATLTCTVSNMNVNADGLEVSW-WKEGKGLLETALGK 187
Qy 276 PREQVNSTYRVSVVTLVHODMLNGKEYCKVSKALPAPIETKISKAK-GQPREQVY 334
Db 188 -RVLQSGTLVVDVAATVVCASEWGGGQYCKVNHPLLPMEKMKKTKASNARPSVY 246
Qy 335 TLPSRDELTKNQ-VSLTCLVKGYPSPDIAYEWESNG--PENNYKTPPVLDSD--DGSF 389
Db 247 VFPPPTQLNGNQRLSTTCMAQGNPRLFRWRKNEPRLPQSSVTSAPMAENPENESTY 306
Qy 390 FLVSKLTVDKSRMOQGNVFCSSVNHIALHNYTKSLSLSPG 431

```

```

Db 307 VAVSVLGVGAEMWAGNVYTCVGVHEALPLQLAQKSVDRASG 348

```

## RESULT 91

Ig mu chain C region - axolotl (fragment)

```

C/Species: Ambystoma mexicanum (axolotl)
C/Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
C/Accession: A46532
R/Fellah, U.S.; Miles, M.V.; Charlemagne, J.; Schwager, J.
Eur. J. Immunol. 22, 2595-2601, 1992
A/Title: Evolution of vertebrate IgM: complete amino acid sequence of the constant region
A/Reference number: A46532; MUID:93011455; PMID:1382992
A/Accession: A46532
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-454 <FEL>
A/Note: sequence inconsistent with the nucleotide translation
A/Note: sequence extracted from NCBI backbone (NCBIP:115354)
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin

```

Query Match 11.0%; Score 296; DB 2; Length 454;

Best Local Similarity 23.1%; Pred. No. 2.2e-10;

Matches 111; Conservative 65; Mismatches 190; Indels 114; Gaps 17;

```

Qy 21 AAQGNKVV---LGKKGDTVELTCTAS--QKSIQHMKNQNIK---LGNGSFLTK 71
Db 1 AAQAPSVPLISCGASLDPVVGCLAKGFLPDSVTFDMDTKNNASFGAVAKLPSTVTTG 60
Qy 72 GPEKLNDRADRSRLMDQGNFPLIKNLKTEDSDTYICEVEDQKEEVQLVFGLTANSPT 131
Db 61 GLVSASRVRVAPBESW-----KNDPYCKV-----T 87
Qy 132 HLIQGSLLTLESPPG--SSPSVQCRSP-----R 159
Db 88 HPSLIGPFLTKVQFVSQVRSKPTVTLHAPAREDINNATVTCIRGFHPQISIKMMKN 147
Qy 160 GKNIOGGKTLVSQLELQDSGTWTCTVLQNKVFEKIDIV-PCPAPB-----PKSCD 211
Db 148 GKXVTS--LYTEPVAADTGNFDPVSLNIEPDMMDTVYSCVVDQTSKFMNTNMS 205
Qy 212 KTHTCPELIGSPVLEPPPKPK-DTLMISRTPEVTCVVDV-----SHEDPEVKF 260
Db 206 KSNLCAQGVPAVTAFTVAPTEDEMESSAVNTCLVTMMGTREGNITMSREDTNEVL 265
Qy 261 NMVYDGEVHNAAKTPREEQYNSTYRVSVVTLVHODMLNGKEYCKVSKALPAPIEKT 320
Db 266 KTEITNPIDFHD-----NATLSVMGIATVCAQDMDAHFKVCVQLHDLAEGRVLS 315
Qy 321 ISKAKG-QPREQVYTLTPSRDELTKNQVSLVLCVKGYPSPDIAYEWESNGO--PENNY 376
Db 316 LQKPNQGRKRKPSVYIYPPSEELALNETVITVCLMKGYPCDLPFRLNLSQQLQKQDY 375
Qy 377 KTPPVLDSD-----SSFLYSKLTVDKSRMOQGNVFCSSVNHIALHNYTKSLSLSPG 431
Db 376 VNTKQAEVDPTTGQSCFPYSLMLKIPAQMTAGNTTTCVGVHEALPLQITQKSIDRSPG 435

```

## RESULT 92

Ig alpha chain C region - rabbit (fragment)

```

C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C/Accession: S09272
R/Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A/Title: The Iga heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A/Reference number: S09264; MUID:90076124; PMID:2512120
A/Accession: S09272
A/Status: not compared with conceptual translation
A/Molecule type: DNA

```

A:Residues: 1-343 <BUR>  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:232-305/Domain: immunoglobulin homology <IMM>

Query Match 10.9%; Score 295.5; DB 2; Length 343;  
Best Local Similarity 31.7%; Pred. No. 1.7e-10;  
Matches 96; Conservative 42; Mismatches 120; Indels 45; Gaps 16;

QY 146 PEGSSSVCCRRPRKGNIGKTLISVQLELDSGVTCTVLQNKQKVEFKIDI-VPCP- 203  
DB 50 PVPSPSSSLYTT-----SSLISLTLDQCPRDGVTCHVKNYDEGQ---DLTPPCOD 99

QY 204 ----APEPKCDKHTTCPELLGSGSVFLFPKPKDITMTSRTPVYCVVVDVSHEDPE-V 258  
DB 100 CHCYCP-PTSC-----GEPLSLQRPDIDGLLESKA-SLTCTISGL--KDPBGA 145

QY 259 KENWY-VDSVEVHNATKPREEOYN-STYRVSVLTVLHQMNGKEYCKVSNKALP-A 315  
DB 146 VFTWEPTNG---NEFVQSVQSVPCGCVSVSLPGCAEPNAGTEFTCTVTHPEIEGG 201

QY 316 PIEKTISSKAKGQREPOVYTLTPPSRDELTKN-QVSLTCLVKGFPYSDIAVEMESNGO--- 371  
DB 202 SLTATISISRGSLTPQVHLPPTEBELANEQVTLTCLVQGRSPKQDLVSWTHNGTLVY 261

QY 372 PENNY---KTPPVLDSDGSPFLYSKLTVDKSRMOQGNVFGSVMEALHNHYTKSL 428  
DB 262 PKDSYLVWKPLPEPGODPTTAVTSLIRVSAEDMNGDSYCVGHEGLAEHPTOKTIDR 321

QY 429 SPG 431  
DB 322 QAG 324

RESULT 93  
Ig alpha-2 chain C region (allotype A2m(1)) - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 16-Jul-1999  
C:Accession: B22360  
R:Flanagan, J.G.; Lefranc, M.P.; Rabbitts, T.H.  
Cell 36, 681-688, 1984  
A:Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1 and  
A:Reference number: A94653; MUID:84130179; PMID:6421489  
A:Accession: B22360  
A:Molecule type: DNA  
A:Residues: 1-340 <FLA>  
C:Genetics:  
A:Gene: GDB:IGHA2  
A:Cross-references: GDB:119333; OMIM:147000  
A:Map position: 14q32.33-14q32.33  
A:Introns: 1/1 103/1; 210/1  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:230-302/Domain: immunoglobulin homology <IMM>

Query Match 10.9%; Score 294.5; DB 2; Length 340;  
Best Local Similarity 29.2%; Pred. No. 1.9e-10;  
Matches 103; Conservative 50; Mismatches 143; Indels 57; Gaps 18;

QY 122 VEGLTNSDTH-----LLOG---QSLTLTLESPPSSPSVQGRS-PRGNIOGKGT 168  
DB 8 VFPLSLDSTPQDGNVYVACLVQGFPPQBPUSVTWSE---SGQVTKRNPSPSDASGDLY 64

QY 169 LSVSQUEL-----QDSGTWCTVLQNKQKVEFKIDI-VPCPAPEPKCDKHTTCPELLG 222  
DB 65 TTSSTQLTLBATQCPDQKSVTCHVHTNSQ---DVTVCVPVPPPPC-----C-----H 111

QY 223 PSVFLPEPKKDTLMTSRTPVYCVVVDVSHEDPEKFWYVDGVEVHNATKPREEOYN 282  
DB 112 PRLSLRPALED-LLGSEANLTKTLTG-LRDASGATFTWPSGK--SAVQGPERRDL 167

QY 283 STYRVSVLTVLHQMNGKEYCKVSNKALPAPIEKTISSKAKGQREPOVYTLTPSRDE 342

DB 168 GCYSVSVLPGCAQPMNGETFTCAAPBELKPLTANITKS-GNTPREVLHLPSPSE 226

QY 343 LTKNQ-VSLTCLVKGFPYSDIAVEMESNGO--PENNYKTPVLD-SG--SFLYSKLT 396  
DB 227 LALNELVLTCLARFSPDVLVRLQGSQELPRKRYLTWASROBPQGTTFVATSLIR 286

QY 397 VDKSMOQGNVSGSVMEALHNHYTKSLSPG-----LQDENC 438  
DB 287 VAAEDWKKGDTFTCNVGHALPLAFTOKTIDRLAKGPHVNVSVMAEVDGIC 339

RESULT 94  
HRKCS  
Ig mu chain C region, secreted (clone 3050) - horn shark  
C:Species: Heterodontus francisci (horn shark)  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Jul-1999  
C:Accession: S01853  
R:Kokubu, F.; Hinde, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.  
EMBO J. 7, 1979-1986, 1988  
A:Title: Complete structure and organization of immunoglobulin heavy chain constant regi  
A:Reference number: S00980; MUID:88328985; PMID:3138109  
A:Accession: S01853  
A:Molecule type: DNA  
A:Residues: 1-438 <KOK>  
A:Cross-references: EMBL:X07781  
A>Note: the sequence was determined from the germline gene  
C:Genetics:  
A:Introns: 100/1; 206/1; 309/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kai  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1;  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob  
F:1-438/Domain: C region <CRE>  
F:20-87/Domain: immunoglobulin homology <IM1>  
F:123-190/Domain: immunoglobulin homology <IM2>  
F:228-291/Domain: immunoglobulin homology <IM3>  
F:330-400/Domain: immunoglobulin homology <IM4>  
F:164,200,245,275,374,411,415,425/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match 10.9%; Score 294; DB 1; Length 438;  
Best Local Similarity 27.1%; Pred. No. 2.8e-10;  
Matches 115; Conservative 65; Mismatches 153; Indels 92; Gaps 21;

QY 53 WNSNQIKLNGQSFLLTKGPEKLNDRADSRSLDQGNFPIIKLTKEDSDTYICEV- 111  
DB 41 WKDNEPITTG---LKTYPVLNKKGTYYTQS-----SQLTISEVSGSKIV-CEVR 88

QY 112 -----EDQKEVQLVFG--LTNSDTHLQGSILTL-- 142

DB 89 RGEVWIKELIPCKDKKXNHPVYILTQSSSEETSRFATVLCSTIDFH---PESTIVML 145

QY 143 -----LESPPSSPSVQGRS-PRGN-IOGKTLISVQLELDSGVTCTVLQNKQKVEF 195  
DB 146 KQGMESGFVSPF--C-----GVNGTFSATSRLLVPAREWNTNRYTQV--SHQGVYQ 197

QY 196 KIDI---VPCPAPEPKCDKHTTCPELLGGSVFLFPKPKDITMTSRTPVYCVVVDV 251  
DB 198 SHNITGQVPC-----SCN-----DEVIKILRPISIEQVL-LKATVLTLCV--V 237

QY 252 SHEDPEKFWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHQMNGKEYCKVSNK 311  
DB 238 SNAAPGVNVSW---TOEQKSLSELAVOGEDADSVISTVNSTQAMLSGAELYCVVNHQ 294

QY 312 ALPAPIEKTISSKAKGQ-REPOV-YTLTPSRDELTKNOVSLTCLVKGFPYSDIAVEMESN 369  
DB 295 DLPTLRASISIHKEEYVDLREPSVSLISPAEDVSAQRFLSLTCLVARGFPREITFIKMTVN 354

QY 370 GQPEN--NYKTPPVLDSD-GSFLYSKLTVDKSMOQGNVSGSVMEALHNHYTKSL 426  
DB 355 DKSVPNGVYKNTENVAMENDNSYFTYSLISIAEBMAGASYSVVGHEALPLKTIINRTV 414

QY 427 SLSPG 431

Db 415 NKSSG 419

## RESULT 95

IG alpha chain C region - rabbit (fragment)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 28-Jan-1985 #sequence\_revision 28-Aug-1985 #text\_change 16-Jul-1999  
 C:Accession: A02174  
 R:Knight, K.L.; Matens, C.L.; Stoklosa, C.M.; Schneiderman, R.D.  
 N:Knight, K.L.; Matens, C.L.; Stoklosa, C.M.; Schneiderman, R.D.  
 A:Title: Genes encoding alpha-heavy chains of rabbit IgA: characterization of cDNA encod  
 A:Reference number: A02174; MUID:84144059; PMID:6322114  
 A:Accession: A02174  
 A:Molecule type: mRNA  
 A:Residues: 1-299 <N1>  
 A:Cross-references: GB:X00353; NID:91575; PID:CAA25100.1; PID:91576  
 C:Comment: This immunoglobulin belongs to the IgA-g subclasses. It was isolated from a rab  
 C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) h  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin; plasma  
 F:86152/Domain: immunoglobulin homology <IM2>  
 F:189-261/Domain: immunoglobulin homology <IM2>  
 F:138-286/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.8%; Score 293; DB 1; Length 299;

Best Local Similarity 31.7%; Pred. No. 28-10; Matches 78; Conservative 42; Mismatches 106; Indels 20; Gaps 10;

Qy 200 VCPAPPEKSCDKHTCP-----ELGG--PSVLPFPKPKDTLMISRTPEVTCVAVDVS 252  
 Db 41 LVPFPDCCPNASCCTCPSSSRNLSGQFSLQRPDLGD-LILGRDASLTCTLSGLK 99  
 Qy 253 HEDPEVKPMYVDGVEVNAKTKPREBOYNSTRVSVLTVLHQMVGKCKVSKA 312  
 Db 100 NPEDEV-FTW--EPTNGNEPVQQRQORDLSGYSVSVLPSSAEKMKARTEFTCTVTPE 156  
 Qy 313 LPA-PIEKTISKAKGQPREPQVYTLPPSRDELTKN-QVSLTCLVKGFPYSDIAVESNG 370  
 Db 157 IDSGSLTATISRGVTP--PQVHLPPPEBELALNEQVTLTCLVGRFSPKDVLSMRHQG 214  
 Qy 371 Q--PENNY---KTPPVLDSDGSFPLYSKLTVDKSRMOQGNVSCVMEALHNHYTQKS 425  
 Db 215 GEVPEPDSFLWKMSPRESSQDKATVAITSLRVPAEDMNGDYSQVGHGLAEHFTQKT 274  
 Qy 426 LSLSPG 431  
 Db 275 IDRLAG 280

## RESULT 96

S09276  
 IG alpha chain C region - rabbit (fragment)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 16-Jul-1999  
 C:Accession: S09276  
 R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.  
 EMBL J. 8, 4041-4047, 1989  
 A:Title: The IgA heavy-chain gene family in rabbit: cloning and sequence analysis of 13  
 A:Reference number: S09264; MUID:90076124; PMID:2512120  
 A:Accession: S09276  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-338 <BUR>  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:228-300/Domain: immunoglobulin homology <IM2>

Query Match 10.8%; Score 292; DB 2; Length 338;

Best Local Similarity 29.2%; Pred. No. 2-7e-10; Matches 92; Conservative 50; Mismatches 133; Indels 40; Gaps 15;

Qy 146 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWC-----TVLQNOCK-----V 193  
 Db 16 PPG--PLVSWTVNGENV-----SAYNFPQAGSGTGPYTAGSELPLPTQCGGQSACHV 70  
 Qy 194 EFKIDI---VCPAPPEKSCDKHTCP-----ELGG--PSVLPFPKPKDTLMISRTPE 243  
 Db 71 EYNSVINBSLVPFPDCCPNASCCTCPSSSRNLSGQFSLQRPDLGD-LILGRDAS 129  
 Qy 244 VTCVAVDVSHEDEPVKPMYVDGVEVNAKTKPREBOYNSTRVSVLTVLHQMVGKCKVSKA 303  
 Db 130 LITCLSGIKNPEDEV-FTW--EPTNGNEPVQQRQORDLSGYSVSVLPSSAEKMKARTE 186  
 Qy 304 YKCVSNKALPA-PIEKTISKAKGQPREPQVYTLPPSRDELTKN-QVSLTCLVKGFPYSD 361  
 Db 187 FCTVTHPEIDSGSLTATISRGVTP--PQVHLPPPEBELALNEQVTLTCLVGRFSPKD 244  
 Qy 362 IAVMESNGQ--PENNY---KTPPVLDSDGSFPLYSKLTVDKSRMOQGNVSCVMEALHNHYTQKS 416  
 Db 245 VLVSWMRQGOEVPEDSFLWKMSPRESSQDKATVAITSLRVPAEDMNGDYSQVGHGLAEHFTQKT 304  
 Qy 417 LHNHYTQKSLSPG 431  
 Db 305 LAEHFTQKTDRLAG 319

## RESULT 97

S09266  
 IG alpha chain C region - rabbit (fragment)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 16-Jul-1999  
 C:Accession: S09266  
 R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.  
 EMBL J. 8, 4041-4047, 1989  
 A:Title: The IgA heavy-chain gene family in rabbit: cloning and sequence analysis of 13  
 A:Reference number: S09264; MUID:90076124; PMID:2512120  
 A:Accession: S09266  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-352 <BUR>  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:241-314/Domain: immunoglobulin homology <IM2>

Query Match 10.8%; Score 292; DB 2; Length 352;

Best Local Similarity 31.2%; Pred. No. 2-8e-10; Matches 95; Conservative 41; Mismatches 118; Indels 50; Gaps 16;

Qy 139 LTTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWCITVLONQKVEFKID 198  
 Db 69 LSLTPEQCP--EDHNVVCRVEH--NYDEGQNLTVLVPCKDPNS-----D 109  
 Qy 199 IVPAPPEKSCDKHTCPBELGGPSVFLPFPKPKDTLMISRTPEVTCVAVDVSHEDE 257  
 Db 110 PTPPCP-PTTC-----GEPSSLQRPDGD-LLESNLSLTCTLSGL--KDEEG 155  
 Qy 258 VKFPMY--VDGVEVNAKTKPREBOYN--STRVSVLTVLHQMVGKCKVSKNKLAP- 314  
 Db 156 AVFTWNPFTNGNEPVQGT---QSYPCGCYSVSVLPGCAEPNAGAEFTCTVTHPEIBG 211  
 Qy 315 APIEKTISKAKGQPREPQVYTLPPSRDELTKN-QVSLTCLVKGFPYSDIAVESNGQ-- 371  
 Db 212 GSLTATISRGSLTP--PQVHLPPPEBELALNALVTLTCLVGRFSPKDVLSMRHNGTLV 269  
 Qy 372 -PENNY---KTPPVLDSDGSFPLYSKLTVDKSRMOQGNVSCVMEALHNHYTQKS 427  
 Db 270 VPKDSFLWKMXPPEPGEPPTTVAITSLRVPAEDMNGDYSQVGHGLAEHFTQRTID 329  
 Qy 428 LSPG 431  
 Db 330 RLAG 333



## RESULT 98

A34891

Ig heavy chain precursor V region - ladyfish

C:Species: Elops saurus (ladyfish)

C:Date: 20-Jul-1990 #sequence\_revision 20-Jul-1990 #ext\_change 21-Jul-2000

C:Accession: A34891

R:Amemiya, C.T.; Litman, G.W.

Proc. Natl. Acad. Sci. U.S.A. 87, 811-815, 1990

A:Title: Complete nucleotide sequence of an immunoglobulin heavy-chain gene and analysis

A:Reference number: A34891; PMID:9018916; PMID:2105490

A:Accession: A34891

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-568 &lt;AME&gt;

A:Cross-references: GB:M26182; NID:g213134; PIDN:AAA49238.1; PID:g213135

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keyword: heterotrimer; immunoglobulin

F:33-116/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 10.7%; Score 290; DB 2; Length 568;

Best Local Similarity 22.3%; Pred. No. 6.5e-10;

Matches 92; Conservative 78; Mismatches 152; Indels 90; Gaps 15;

QY 65 OGSFLLTKGPSKLNDRADSRSLMDGNFLLI-----KNLKID 103

DB 207 EGDFFSKAEKPKCTTELGKKT-----PVIKPPPKPPRPQPVLSIMTPSGELTLNK 259

QY 104 SDTYTEVEDQKEVQLVFLGLTANSDTHLLOGLTLLESPPSSPSVQCRSPRKNI 163

DB 260 TATFALCLAND-----FYPKGHSFKMLRDCEVT-----DGIATLTCECKKDKSF 304

QY 164 OGSKTLVSQLELD--SGMTCTVLNOQKVE--FKIDVPCAPAPKSGDKHTHPRL 220

DB 305 TASSFLQASBSQKRLDGFITCFDGETTEGTIVKSAEC--SPQAQIDAK----- 355

QY 221 GGSFVLPFPKPKDTLMISRTPEVTCVVDVSHDEPVFNMYVDGE--VHNAKTKPRE 278

DB 356 -----ISPPTPEE-LFLQGTTRLTC-----KITGVVDVNRVNTMEVSGEVAV 396

QY 279 EQYNSTYRVVSVLTVLHODWLNGEKYCKVSKALPAPLEKITSK-AKQOPREPVYTLR 337

DB 397 GGFDE-QKMISKLLIYEEKNRTEYTCVKEHSDLPSTPSYRRCGGKMGSPFVFLA 455

QY 338 PSRDELTKQVSLTCLVKGFYPSDIAVEMESNGEPENNTKTPPVLDSDGSFELYSLTV 397

DB 456 PA-EQNNLSIVILICAKDYPFQVILSWLDQPEYTDVPTTEVVKTEYSVFSQLTI 514

QY 398 DKSMMOQGVFSCSVNHEAL-----HNHYTKSLSPGLQDLDTTC 438

DB 515 PASDMVSGVYVSCAVHETVMEVSVKTIYRTDVSVKKPTTVSLDLNVPQTC 566

## RESULT 99

S09265

Ig alpha chain C region - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #ext\_change 16-Jul-1999

C:Accession: S09265

R:Burrett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.

EMBO J. 8, 4041-4047, 1989

A:Title: The IgA heavy-chain gene family in rabbit: cloning and sequence analysis of 13

A:Reference number: S09264; PMID:90076124; PMID:2512120

A:Accession: S09265

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-357 &lt;BUR&gt;

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keyword: immunoglobulin

F:246-319/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 10.7%; Score 289; DB 2; Length 357;

Best Local Similarity 28.2%; Pred. No. 4.3e-10;

Matches 93; Conservative 49; Mismatches 122; Indels 66; Gaps 14;

QY 134 LOGQSILTL-TLESPPGS-----SPSVOCR---SPRKNIOGKTLISVS 172

DB 43 VSGSVIPEPPVSPSSSLYTTYSILRLPAEQCPRENSVACRVBNHKKQDVVPPSPACH 102

QY 173 QLELDGSGTMTCTVLONQKVEFKIDIVPCAPAPKSCDKHTCEBLGGSFVLPKPK 232

DB 103 ESTIEPPTPTC-----PCPCPSP-SC-----GKPLSLQRPDL 135

QY 233 KDTLMISRTPEVTCVVDVSHDEPVKFNWYDGEVHNAKTKPRE---QNSTYRVVS 289

DB 136 GD-LILNSNASLTCRLKGLNPEGAV-FTW-----EPFGKSPVQOSPOLDGCYSVS 188

QY 290 VLTVLHODWLNGEKYCKVSKNKAIPA-PEKITSKAKGQPRBPVYTLPPSRDELTKNQ 347

DB 189 VLPGCATVNNAGTEFTCTVTHPEIBODSLTGITSDKDSLRPQVHLPPSEBLALNLY 248

QY 348 VSLTCLVKGFPYPSDIAVEMESNGQ---PENNYKTPPVLD---SDGSFVLSKLTVDKSR 401

DB 249 VTLTCLVGRFSPKDVLSVTHNGTVPVPRKDSVLWKPLREPGDPTTVAITSLRVPAD 308

QY 402 MOQGVNFCSCSVNHEALHNHYTKSLSPG 431

DB 309 KNOGDSYSCVGVGHEGLAEHFTOKTIDRLAG 338

## RESULT 100

I56230

Ig alpha-2 chain - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #ext\_change 21-Jan-2000

C:Accession: I56230

R:Chitalacharuvu, K.R.; Raines, M.; Morrison, S.L.

J. Immunol. 152, 5299-5304, 1994

A:Title: Divergence of human alpha-chain constant region gene sequences. A novel recomb.

A:Reference number: I56230; PMID:94246170; PMID:8189047

A:Accession: I56230

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-340 &lt;RES&gt;

A:Cross-references: GB:S71043; NID:9546798; PIDN:AA830803.1; PID:9546799

C:Gene(s):

A:Gene: IGA2

A:Introns: 103/1; 210/1

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:330-302/Domain: immunoglobulin homology &lt;IMM&gt;

Query Match 10.6%; Score 286.5; DB 2; Length 340;

Best Local Similarity 28.9%; Pred. No. 5.7e-10;

Matches 102; Conservative 50; Mismatches 144; Indels 57; Gaps 18;

QY 122 VFGLTANSDTH-----LLQG---OSLTLLESPPSSPSVQCRS-PRGKNIOGKKT 168

DB 8 VFPLSDSTPQGVNVVACTVGGFPQEPPLSVTWSB---SGQNTVARNPPQDASGDLY 64

QY 169 LVSQLEL-----QDSGTCTVLONQKVEFKIDI-VPCAPAPKSCDKHTCEBLDG 222

DB 65 TTSQQLTPPATQCPGKSTVCHVKHYTNSQ---DVTPVCRVPPPPC-----C-----H 111

QY 223 PSVFLPFPKPKDTLMISRTPEVTCVVDVSHDEPVKFNWYDGEVHNAKTKPREQYN 282

DB 112 PRLSLHRLPDEL-LILGSANILCTLTGL-RASGATFTWTSBGK--SAVGGPEBRDLC 167

QY 283 STYRVSVLTVLHODWLNGEKYCKVSKNKAIPAPEKITSKAKGQPRBPVYTLPPSRDE 342

DB 168 GCVSVSVLPVPGCAQPMNHGTEFTCTAARHKLTPITANITKS-GMTFREVHLPPSE 226

QY 343 LTKNQ-VSLTCLVKGFPYPSDIAVEMESNGQ---PENNYKTPPVLD-SDG--SFVLSKLT 396

DB 227 LALNELVTLTCLARGSPKDVLRWLQGSQELPRKRYLTMASROBPQGTTFATVTSILR 286

QY 397 VKSMMOQGVNFCSCSVNHEALHNHYTKSLSPG-----LQDLDTTC 438

Db 287 VAEDMKKGDFTSCMVGHEALPLAFTOKTIDRLAGKPTHVNSVYMAEVDGTC 339

Search completed: August 3, 2004, 13:15:19  
Job time : 16.7668 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: August 3, 2004, 13:01:34 ; Search time 19.0695 Seconds  
(without alignments)  
1754.300 Million cell updates/sec

Title: SEQ07  
Perfect score: 3414  
Sequence: 1 MNRGVPRHLLVLQLALP.....DETCAEAQDGLDGLWTTDP 648

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 369414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 369414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 125 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/6C.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

102(12) SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3189	93.4	630	4	US-08-472-888A-6
2	2116	62.0	530	3	US-08-477-460B-4
3	2116	62.0	530	3	US-08-379-516-4
4	2116	62.0	530	3	US-09-329-916-4
5	2116	62.0	530	3	US-08-485-372A-4
6	2116	62.0	530	4	US-09-409-006A-4
7	2116	62.0	530	4	US-08-484-681-4
8	2116	62.0	530	5	PCT-US93-07422-4
9	2059	60.3	432	3	US-08-477-460B-2
10	2059	60.3	432	3	US-08-379-516-2
11	2059	60.3	432	3	US-09-329-916-2
12	2059	60.3	432	3	US-08-485-372A-2
13	2059	60.3	432	4	US-09-409-006A-2
14	2059	60.3	432	4	US-08-484-681-2
15	2059	60.3	432	5	PCT-US93-07422-2
16	2036.5	59.7	532	2	US-08-417-485-6
17	2036.5	59.7	532	2	US-08-284-391B-6
18	2036.5	59.7	532	3	US-09-218-950-6
19	2036.5	59.7	532	5	PCT-US92-01785-6
20	2036.5	59.7	532	5	PCT-US95-00454-6
21	2036.5	59.7	575	2	US-08-417-495-4
22	2036.5	59.7	575	2	US-08-284-391B-4
23	2036.5	59.7	575	3	US-09-218-950-4
24	2036.5	59.7	575	5	PCT-US92-01785-4
25	2036.5	59.7	575	5	PCT-US95-00454-4
26	2032.5	59.5	462	2	US-08-417-495-5
27	2032.5	59.5	462	2	US-08-284-391B-5

28	2032.5	59.5	462	3	US-09-218-950-5	Sequence 5, Appli
29	2032.5	59.5	462	5	PCT-US92-01785-5	Sequence 5, Appli
30	2032.5	59.5	462	5	PCT-US95-00454-5	Sequence 5, Appli
31	2029	59.4	398	2	US-08-284-391B-29	Sequence 29, Appli
32	2029	59.4	398	2	US-09-218-950-29	Sequence 29, Appli
33	2029	59.4	458	3	US-08-466-368-4	Sequence 4, Appli
34	2021	59.2	457	4	US-08-328-500-9	Sequence 9, Appli
35	2017	59.1	402	1	US-08-236-311-1	Sequence 1, Appli
36	2017	59.1	402	3	US-08-457-918-1	Sequence 1, Appli
37	2015	59.0	458	3	US-09-039-555B-15	Sequence 15, Appli
38	2007	58.8	458	3	US-09-517-605-3	Sequence 3, Appli
39	2001	58.6	394	3	US-08-466-368-2	Sequence 2, Appli
40	2001	58.6	394	4	US-08-328-500-2	Sequence 2, Appli
41	1989	58.3	458	6	5223394-7	Patent No. 5223394
42	1951	57.1	394	6	522318-2	Patent No. 522318
43	1904	55.8	434	1	US-08-226-311-4	Sequence 4, Appli
44	1904	55.8	434	3	US-08-457-918-4	Sequence 4, Appli
45	1892	55.4	433	2	US-08-867-149-1	Sequence 1, Appli
46	1892	55.4	433	3	US-08-808-374-1	Sequence 1, Appli
47	1887	55.3	433	3	US-09-100-409A-1	Sequence 1, Appli
48	1695	49.6	433	6	517838-13	Patent No. 517838
49	1585.5	46.4	410	3	US-08-630-172-17	Sequence 17, Appli
50	1585.5	46.4	410	3	US-09-375-419-17	Sequence 17, Appli
51	1368	40.1	318	6	5223394-11	Patent No. 5223394
52	1363	39.9	295	6	5223394-9	Patent No. 5223394
53	1338.5	39.2	254	2	US-08-284-391B-33	Sequence 33, Appli
54	1338.5	39.2	254	3	US-09-218-950-33	Sequence 33, Appli
55	1275	37.3	592	4	US-09-313-942-8	Sequence 8, Appli
56	1243	36.4	622	4	US-09-499-846-2	Sequence 2, Appli
57	1237	36.2	704	4	US-09-590-656-2	Sequence 2, Appli
58	1237	36.2	704	4	US-09-733-764-2	Sequence 2, Appli
59	1235.5	36.2	497	4	US-09-499-846-6	Sequence 6, Appli
60	1235.5	36.2	525	4	US-09-499-846-4	Sequence 4, Appli
61	1232.5	36.1	424	5	PCT-US95-03866-12	Sequence 12, Appli
62	1232.5	36.1	424	5	PCT-US95-03866-14	Sequence 14, Appli
63	1232.5	36.1	459	1	US-08-157-101A-7	Sequence 7, Appli
64	1230.5	36.0	446	3	US-08-397-411-7	Sequence 7, Appli
65	1225.5	35.9	488	3	US-08-776-511-2	Sequence 2, Appli
66	1225	35.9	488	4	US-09-499-846-12	Sequence 12, Appli
67	1223.5	35.8	680	3	US-08-227-996C-15	Sequence 15, Appli
68	1222.5	35.8	454	2	US-07-934-373C-22	Sequence 22, Appli
69	1222.5	35.8	454	3	US-08-437-642B-22	Sequence 22, Appli
70	1222.5	35.8	454	4	US-08-146-206C-22	Sequence 22, Appli
71	1222.5	35.8	454	4	US-09-705-686-22	Sequence 22, Appli
72	1222.5	35.8	454	5	PCT-US93-07832-22	Sequence 22, Appli
73	1221.5	35.8	475	4	US-09-740-002-25	Sequence 25, Appli
74	1221.5	35.8	473	3	US-09-049-672A-4	Sequence 4, Appli
75	1220.5	35.7	387	1	US-08-470-299-4	Sequence 4, Appli
76	1220.5	35.7	453	4	US-09-532-856-6	Sequence 6, Appli
77	1220.5	35.7	453	4	US-09-524-100C-6	Sequence 6, Appli
78	1220.5	35.7	497	4	US-09-499-846-10	Sequence 10, Appli
79	1220.5	35.7	525	4	US-09-499-846-8	Sequence 8, Appli
80	1220	35.7	388	3	US-09-131-247-16	Sequence 16, Appli
81	1219.5	35.7	453	4	US-09-301-593-18	Sequence 18, Appli
82	1219.5	35.7	475	4	US-09-740-002-27	Sequence 27, Appli
83	1219.5	35.7	475	4	US-08-487-550-12	Sequence 12, Appli
84	1219.5	35.7	476	4	US-09-526-098-12	Sequence 12, Appli
85	1216	35.6	442	4	US-08-472-888A-7	Sequence 7, Appli
86	1216	35.6	442	5	PCT-US96-10043-9	Sequence 9, Appli
87	1215.5	35.6	437	3	PCT-US96-10043-11	Sequence 11, Appli
88	1215	35.6	476	2	US-08-378-939-11	Sequence 10, Appli
89	1213.5	35.5	446	4	US-09-157-452B-12	Sequence 12, Appli
90	1213	35.5	449	1	US-08-458-516-13	Sequence 13, Appli
91	1212.5	35.5	468	4	US-09-485-378B-67	Sequence 67, Appli
92	1212.5	35.5	711	4	US-09-485-378B-90	Sequence 90, Appli
93	1212.5	35.5	859	4	US-09-313-942-7	Sequence 7, Appli
94	1212.5	35.5	951	4	US-09-313-942-9	Sequence 9, Appli
95	1212	35.5	451	4	US-09-247-352-3	Sequence 3, Appli
96	1212	35.5	451	4	US-09-466-635-3	Sequence 3, Appli
97	1212	35.5	472	4	US-09-301-593-10	Sequence 30, Appli
98	1211.5	35.5	232	2	US-08-595-043A-50	Sequence 50, Appli
99	1211.5	35.5	331	3	US-09-178-669-2	Sequence 2, Appli
100	1211.5	35.5	331	4	US-09-761-413-2	Sequence 2, Appli

101 1211.5 35.5 360 4 US-09-180-100-11 Sequence 11, App1  
 102 1211.5 35.5 371 1 US-08-236-311-7 Sequence 7, App1  
 103 1211.5 35.5 371 3 US-08-457-918-7 Sequence 7, App1  
 104 1211.5 35.5 376 4 US-09-180-100-22 Sequence 22, App1  
 105 1211.5 35.5 366 2 US-08-784-512-3 Sequence 3, App1  
 106 1211.5 35.5 366 3 US-09-176-228-3 Sequence 3, App1  
 107 1211.5 35.5 452 3 US-09-027-449-71 Sequence 71, App1  
 108 1211.5 35.5 452 3 US-09-026-985-71 Sequence 71, App1  
 109 1211.5 35.5 452 4 US-09-121-952A-71 Sequence 71, App1  
 110 1211.5 35.5 452 4 US-09-234-340A-71 Sequence 71, App1  
 111 1211.5 35.5 476 3 US-08-487-550-4 Sequence 4, App1  
 112 1211.5 35.5 476 4 US-09-526-098-4 Sequence 4, App1  
 113 1211.5 35.5 478 3 US-08-487-550-8 Sequence 8, App1  
 114 1211.5 35.5 478 4 US-09-526-098-8 Sequence 8, App1  
 115 1211.5 35.5 547 4 US-09-746-359A-54 Sequence 54, App1  
 116 1211.5 35.5 571 4 US-09-746-359A-53 Sequence 53, App1  
 117 1211 1211 450 4 US-09-532-856-5 Sequence 5, App1  
 118 1211 1211 450 4 US-09-524-100C-5 Sequence 5, App1  
 119 1211 1211 472 4 US-09-301-593-43 Sequence 43, App1  
 120 1210 35.4 451 2 US-08-887-352B-18 Sequence 18, App1  
 121 1210 35.4 451 3 US-09-109-207C-18 Sequence 18, App1  
 122 1210 35.4 451 3 US-09-282-505-2 Sequence 2, App1  
 123 1210 35.4 451 3 US-09-054-255-2 Sequence 2, App1  
 124 1210 35.4 451 3 US-09-296-005-18 Sequence 18, App1  
 125 1210 35.4 451 4 US-09-282-846-2 Sequence 2, App1

## ALIGNMENTS

RESULT 1  
 US-08-472-888A-6  
 ; Sequence 6, Application US/08472888A  
 ; Patent No. 6613746

GENERAL INFORMATION:  
 APPLICANT: Seed, Brian  
 APPLICANT: Walz, Gerald  
 TITLE OF INVENTION: AGP-ANTIBODY FUSION PROTEINS  
 TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Clark & Elbing LLP  
 STREET: 176 Federal Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/472,888A  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 07/618,314  
 FILING DATE: 23-NOV-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elbing, Karen L.  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/258001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-428-0200  
 TELEFAX: 617-428-7045  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 630 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: linear

MOLECULE TYPE: protein  
 US-08-472-888A-6  
 Query Match 93.4%; Score 3189; DB 4; Length 630;  
 Best Local Similarity 97.8%; Pred. No. 5,1e-245;  
 Matches 617; Conservative 0; Mismatches 6; Indels 8; Gaps 4;

1 NMRGVPFRLLVLTALIPATQGNKVYLGKGGPTVELTCASQKSIQFMKSNQIK 60  
 1 NMRGVPFRLLVLTALIPATQGNKVYLGKGGPTVELTCASQKSIQFMKSNQIK 60  
 61 IINGGSPFTTKPSKLNDRASRSRSLMDQGNPILIIKNKIDSDTYICEVEDQKEVLT 120  
 61 IINGGSPFTTKPSKLNDRASRSRSLMDQGNPILIIKNKIDSDTYICEVEDQKEVLT 120  
 121 LVFGTLTNSDTHLLOQGSITLTLESPPGSSPSVQCRSPGKNIQGGKTLVSQLEIQDSG 180  
 121 LVFGTLTNSDTHLLOQGSITLTLESPPGSSPSVQCRSPGKNIQGGKTLVSQLEIQDSG 180  
 181 TWCTVLTQNKQKVEERKIDIVLAFOKASIVYKKEGEVFSFPLAFTVEKLTGSGELMW 240  
 181 TWCTVLTQNKQKVEERKIDIVLAFOKASIVYKKEGEVFSFPLAFTVEKLTGSGELMW 240  
 241 QMRASSSKSWITPDLKKNKESVKEVTQDPKIQMGKULPLHLTLPOALPOVAGSGNLTLLA 300  
 241 QMRASSSKSWITPDLKKNKESVKEVTQDPKIQMGKULPLHLTLPOALPOVAGSGNLTLLA 300  
 301 LEAKTGKLGHOEVNLYVMRATQLOKNLTCEVWGPTSPKMLSLKENKAKVSKREKPVW 360  
 301 LEAKTGKLGHOEVNLYVMRATQLOKNLTCEVWGPTSPKMLSLKENKAKVSKREKPVW 360  
 361 LNPBAGMOCCLSDSGQVLEESNIVLPTWSTPVHADPGEPSCKHTCPPCPAPELT 420  
 361 LNPBAGMOCCLSDSGQVLEESNIVLPTWSTPVHADPGEPSCKHTCPPCPAPELT 420  
 415 GGPVSFLPPPKKDTLMTSRTPEVTCVVDVSHEDPEVFNMYVDGVEYHNAKTKRREO 474  
 421 GGPVSFLPPPKKDTLMTSRTPEVTCVVDVSHEDPEVFNMYVDGVEYHNAKTKRREO 479  
 475 YNSTYRVSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSR 534  
 480 YNSTYRVSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSR 538  
 535 DELTKNOVSLTCLVNGFVPSDIAVEMESNGOPENNYKTPVLDSGSEFFYLSKLTVDKS 594  
 539 DELTKNOVSLTCLVNGFVPSDIAVEMESNGOPENNYKTPVLDSGSEFFYLSKLTVDKS 598  
 595 RMQGNVFSQVMEALHNHYTQKSLSPG 625  
 599 RMQGNVFSQVMEALHNHYTQKSLSPG 629

RESULT 2  
 US-08-477-460B-4  
 ; Sequence 4, Application US/08477460B  
 ; Patent No. 6034223  
 GENERAL INFORMATION:  
 APPLICANT: Progenics Pharmaceuticals, Inc.  
 TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED  
 TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Cooper & Dunham  
 STREET: 30 Rockefeller Plaza  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10112  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.24

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,460B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/927,931  
FILING DATE: 07-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/ATM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 977-9809  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: homo sapien  
CELL TYPE: lymphocyte  
US-08-477-460B-4

Query Match 62.0%; Score 2116; DB 3; Length 530;  
Best Local Similarity 68.9%; Pred. No. 7.3e-160;  
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```
QY 1 MNRGVPFRHLVLTQLALPPATOGKRVVLGKGGDTVELCTASOKSIOPHMKNNOIK 60
DB 1 MNRGVPFRHLVLTQLALPPATOGKRVVLGKGGDTVELCTASOKSIOPHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
QY 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
QY 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKCLPLHLTPQALPYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKCLPLHLTPQALPYAG--SGNL 297
QY 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNGALTSQVH 255
DB 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNGALTSQVH 255
QY 298 TLAEAKTGKLDHENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKEAVSRREK 356
DB 298 TLAEAKTGKLDHENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKEAVSRREK 356
QY 256 TFPAVLQSSGLYSLSVTVPSNFGTQYTCNV-----DHK 292
DB 256 TFPAVLQSSGLYSLSVTVPSNFGTQYTCNV-----DHK 292
QY 357 PAVVLNBPAGMWQCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDHTCP 411
DB 357 PAVVLNBPAGMWQCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDHTCP 411
QY 293 P-----SNTKVDKTVERKCCVCEPCBPAP----- 317
DB 293 P-----SNTKVDKTVERKCCVCEPCBPAP----- 317
QY 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 471
DB 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 471
QY 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 375
DB 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 375
QY 472 EEOYSTYRVVSVLTVLHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
DB 472 EEOYSTYRVVSVLTVLHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
QY 376 EEOGNSYFRVVSVLTVVHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
DB 376 EEOGNSYFRVVSVLTVVHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
QY 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
DB 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
QY 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
DB 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
QY 592 DKSRMOQGNVSCVMHEALHNHYTOKSLSPG 625
DB 592 DKSRMOQGNVSCVMHEALHNHYTOKSLSPG 625
```

Db 496 DKSRMOQGNVSCVMHEALHNHYTOKSLSPG 529  
RESULT 3  
US-08-379-516-4  
Sequence 4, Application US/08379516  
Patent No. 6083478  
GENERAL INFORMATION:  
APPLICANT: Allaway, Graham P.  
TITLE OF INVENTION: No. 6083478-Peptidyl Moiety-Conjugated CD4-Gamma2 and CD4-IgG2  
FILE REFERENCE: 41215-A-PCT-US  
CURRENT APPLICATION NUMBER: US/08/379,516  
EARLIER FILING DATE: 1996-06-10  
EARLIER APPLICATION NUMBER: PCT/US93/07422  
EARLIER FILING DATE: 1993-08-06  
EARLIER APPLICATION NUMBER: 07/927,931  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 530  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-379-516-4

Query Match 62.0%; Score 2116; DB 3; Length 530;  
Best Local Similarity 68.9%; Pred. No. 7.3e-160;  
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```
QY 1 MNRGVPFRHLVLTQLALPPATOGKRVVLGKGGDTVELCTASOKSIOPHMKNNOIK 60
DB 1 MNRGVPFRHLVLTQLALPPATOGKRVVLGKGGDTVELCTASOKSIOPHMKNNOIK 60
QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
QY 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLTQONKKVEFKIDIVLAFOKASSIYKKKEGEOVESPLAFTVEKLTGSGELMW 240
QY 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKCLPLHLTPQALPYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVKRVTPQPKLQMGKCLPLHLTPQALPYAG--SGNL 297
QY 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNGALTSQVH 255
DB 217 -----PCSRSTSESTALGCLVKDYFPEPVTVSNMNGALTSQVH 255
QY 298 TLAEAKTGKLDHENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKEAVSRREK 356
DB 298 TLAEAKTGKLDHENVLVNMRATOL-QKNLTCEVWGPTSPRLMLSLKENKEAVSRREK 356
QY 256 TFPAVLQSSGLYSLSVTVPSNFGTQYTCNV-----DHK 292
DB 256 TFPAVLQSSGLYSLSVTVPSNFGTQYTCNV-----DHK 292
QY 357 PAVVLNBPAGMWQCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDHTCP 411
DB 357 PAVVLNBPAGMWQCLSDSGVLESNIKVLPTWSTPV-----PCPAPBPCKDHTCP 411
QY 293 P-----SNTKVDKTVERKCCVCEPCBPAP----- 317
DB 293 P-----SNTKVDKTVERKCCVCEPCBPAP----- 317
QY 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 471
DB 412 ELGSPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 471
QY 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 375
DB 318 --VAGPSVFLPPPKDPTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVGVHNAKTKPR 375
QY 472 EEOYSTYRVVSVLTVLHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
DB 472 EEOYSTYRVVSVLTVLHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 531
QY 376 EEOGNSYFRVVSVLTVVHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
DB 376 EEOGNSYFRVVSVLTVVHODMLNGEKYKCVSNKALPAPIEKTSKAGQPREQVYTLR 435
QY 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
DB 532 PSRELTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 591
QY 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
DB 436 PSREMTKNOVSLTCLVKGFFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTV 495
QY 592 DKSRMOQGNVSCVMHEALHNHYTOKSLSPG 625
DB 592 DKSRMOQGNVSCVMHEALHNHYTOKSLSPG 625
```

Db 496 DKSRMOOGNVFSCSVMHGALHNHYTOKSLSPG 529

RESULT 4  
US-09-329-916-4  
Sequence 4, Application US/09329916  
Patent No. 6177549

GENERAL INFORMATION:  
APPLICANT: Progenics Pharmaceuticals, Inc.  
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/329,916  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/477,460  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 07/927,931  
FILING DATE: 07-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 977-9809  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: homo sapien  
CELL TYPE: lymphocyte  
US-09-329-916-4

Query Match 62.0%; Score 2116; DB 3; Length 530;  
Best Local Similarity 68.9%; Pred. No. 7,3e-160;  
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

QY 1 MNRGVPFRHLILVQLALPAATQGNKVLGKGDIVELTCTASQKSIQFHWKNSQIK 60  
Db 1 MNRGVPFRHLILVQLALPAATQGNKVLGKGDIVELTCTASQKSIQFHWKNSQIK 60

QY 61 ILNNGSFLTKGPKSLNDRADSRSLMDQNFILITNKLITESPDTYICVEPOKKEVOL 120  
Db 61 ILNNGSFLTKGPKSLNDRADSRSLMDQNFILITNKLITESPDTYICVEPOKKEVOL 120

QY 121 LVFGLTANSDTHLLQGSGLTLESPPGSSPSVQCRSPKNTIQGKTLISVQLELQDSG 180  
Db 121 LVFGLTANSDTHLLQGSGLTLESPPGSSPSVQCRSPKNTIQGKTLISVQLELQDSG 180

QY 181 TWTCVTIQQNQKVEFKIDIVLAFQKASSIVYKKEGSEQVEPSFPLATVEKLTGSGELMW 240  
Db 181 TWTCVTIQQNQKVEFKIDIVLAF--AST-----KGPSV---FPLA----- 216

QY 241 QAERASSKSWITFDLKNKEVSRYKVTODPKLQMGKLLPLHLTPQALPQYAG---SGNL 297  
Db 217 -----PCSRSTSESTALGCLVXDYFPEPTVTSWNSGALTSQVH 255

QY 298 TLAEAKTGKHOENVLYVMRATQL-QKNLTCEVWGPSTPKMLSLKENKAKVSKREK 356  
Db 256 TPEPAVLQSSGLYSLSVTVTPSSNFGTQTYTCNV-----DHK 292

QY 357 FVWVNLPEAGMOCCLSDSGQVLLBSNIXVLPWTSTPV-----PCPAPEKSCDKHTTCTP 411  
Db 293 P-----SNTKVDKVERKCCVCEPCPAP----- 317

QY 412 ELLGSPVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAKTKPR 471  
Db 318 --VAGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVQNMVYDGEVHNAKTKPR 375

QY 472 BEQNSTFRVSVLTVVHODMLNGKEYKCKVSNKALPAPIEKTISKTKGPRREPQVYTLR 531  
Db 376 BEQNSTFRVSVLTVVHODMLNGKEYKCKVSNKALPAPIEKTISKTKGPRREPQVYTLR 435

QY 532 PSRDELTKQVSLTCLVKGFPSPDIWVESNQPENNYKTTTPVLDSSGFFLYSKLTV 591  
Db 436 PSRDEMTKQVSLTCLVKGFPSPDIWVESNQPENNYKTTTPVLDSSGFFLYSKLTV 495

QY 592 DKSRMOOGNVFSCSVMHGALHNHYTOKSLSPG 625  
Db 496 DKSRMOOGNVFSCSVMHGALHNHYTOKSLSPG 529

RESULT 5  
US-08-485-372A-4  
Sequence 4, Application US/08485372A  
Patent No. 6187748

GENERAL INFORMATION:  
APPLICANT: Beaudry, Gary A.  
TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,372A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,227  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 37690-II-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:

ORGANISM: homo sapien  
CELL TYPE: lymphocyte  
US-08-485-372A-4

Query Match 62.0%; Score 2116; DB 3; Length 530;

Best Local Similarity 68.9%; Pred. No. 7.3e-160; Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

QY 1 MNRGVPFRLHLLVQLALLPAATQGNKVLLGKGGDTVELTCTASOKSIOFHMNSNOIK 60
DB 1 MNRGVPFRLHLLVQLALLPAATQGNKVLLGKGGDTVELTCTASOKSIOFHMNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLATVEKLTSGGELMW 240
DB 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLATVEKLTSGGELMW 240
QY 241 QAERASSKSWITFDLKNKEVSVKRVTDPRKLOMGKKLPLHLTLPOALPOYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVKRVTDPRKLOMGKKLPLHLTLPOALPOYAG--SGNL 297
QY 217 -----PCSRSTSESTRALGCLVKDYFPEPVTVSMNSGALTSGVH 255
DB 217 -----PCSRSTSESTRALGCLVKDYFPEPVTVSMNSGALTSGVH 255
QY 298 TLAEAKTGKLGHOEVLNVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 298 TLAEAKTGKLGHOEVLNVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
QY 256 TFPAVLQSSGLYSLSVTVTPSSNFGTQYTCNV-----DHK 292
DB 256 TFPAVLQSSGLYSLSVTVTPSSNFGTQYTCNV-----DHK 292
QY 357 PWWLNPBAGMWQCLSDSGVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
DB 357 PWWLNPBAGMWQCLSDSGVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
QY 293 P-----SNTKVDKTVERKCCVCPCPAP----- 317
DB 293 P-----SNTKVDKTVERKCCVCPCPAP----- 317
QY 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
DB 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
QY 318 --VAGPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 375
DB 318 --VAGPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 375
QY 472 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 531
DB 472 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 531
QY 376 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 435
DB 376 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 435
QY 532 PSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 591
DB 532 PSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 591
QY 436 PSREMTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 495
DB 436 PSREMTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 495
QY 592 DKSRMOQGNVPSGVMEHALNHNTOKSLSPG 625
DB 592 DKSRMOQGNVPSGVMEHALNHNTOKSLSPG 625

```

# RESULT 6

US-09-409-006A-4

Sequence 4, Application US/09409006A

Patent No. 6342586

GENERAL INFORMATION:

APPLICANT: Progenics Pharmaceuticals, Inc.

TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED

TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IGG2 IMMUNOCONJUGATES, AND USES THEREOF

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/409,006A  
FILING DATE: 29-SEP-1999  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/927,931  
FILING DATE: 07-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 977-9809  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: homo sapien  
CELL TYPE: lymphocyte  
US-09-409-006A-4

Query Match 62.0%; Score 2116; DB 4; Length 530;

Best Local Similarity 68.9%; Pred. No. 7.3e-160; Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

QY 1 MNRGVPFRLHLLVQLALLPAATQGNKVLLGKGGDTVELTCTASOKSIOFHMNSNOIK 60
DB 1 MNRGVPFRLHLLVQLALLPAATQGNKVLLGKGGDTVELTCTASOKSIOFHMNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLATVEKLTSGGELMW 240
DB 181 TWCTCTVLOKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLATVEKLTSGGELMW 240
QY 241 QAERASSKSWITFDLKNKEVSVKRVTDPRKLOMGKKLPLHLTLPOALPOYAG--SGNL 297
DB 241 QAERASSKSWITFDLKNKEVSVKRVTDPRKLOMGKKLPLHLTLPOALPOYAG--SGNL 297
QY 217 -----PCSRSTSESTRALGCLVKDYFPEPVTVSMNSGALTSGVH 255
DB 217 -----PCSRSTSESTRALGCLVKDYFPEPVTVSMNSGALTSGVH 255
QY 298 TLAEAKTGKLGHOEVLNVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 298 TLAEAKTGKLGHOEVLNVMRATOL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
QY 256 TFPAVLQSSGLYSLSVTVTPSSNFGTQYTCNV-----DHK 292
DB 256 TFPAVLQSSGLYSLSVTVTPSSNFGTQYTCNV-----DHK 292
QY 357 PWWLNPBAGMWQCLSDSGVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
DB 357 PWWLNPBAGMWQCLSDSGVLLSNIKYLPTWSTPV-----PCPAPEKSCDKHTCP 411
QY 293 P-----SNTKVDKTVERKCCVCPCPAP----- 317
DB 293 P-----SNTKVDKTVERKCCVCPCPAP----- 317
QY 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
DB 412 ELGSPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 471
QY 318 --VAGPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 375
DB 318 --VAGPSVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFMWYDGVENVNAKTKPR 375
QY 472 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 531
DB 472 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 531
QY 376 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 435
DB 376 EEOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISSKAGQPREPOVYTL 435
QY 532 PSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 591
DB 532 PSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 591
QY 436 PSREMTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 495
DB 436 PSREMTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLT 495
QY 592 DKSRMOQGNVPSGVMEHALNHNTOKSLSPG 625
DB 592 DKSRMOQGNVPSGVMEHALNHNTOKSLSPG 625

```



Db 496 DKSROQGNVFCSCVWHEALHNHYTQKSLSPG 529

RESULT 7  
US-08-484-681-4  
Sequence 4, Application US/08484681  
Patent No. 6451313

GENERAL INFORMATION:

APPLICANT: Beaudry, Gary A.

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: CD4-GAMMA2 CD4-IGS2 CHIMERAS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,681

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 37690-II-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 530 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ORIGINAL SOURCE: cDNA

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-08-484-681-4

Query Match 62.0%; Score 2116; DB 4; Length 530;

Best Local Similarity 68.9%; Pred. No. 7.3e-160; Indels 114; Gaps 10;

Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

Db 1 MNRGVPRHLLLVLTALPAPATQGNKVLGKGDVTELTCTASOKKSIQPHNKNNOIK 60

QY 298 TLAEKTKGLHGVNLYVMRATOL-QKULTCVWGPPTPKLMLSLKENKAVSKREK 356

Db 256 TEPVAVQSSGLYSLSVTVTPSSNFGOTQYTCNV-----DHK 292

QY 357 PVWVLPKAGMWOCLISDSGQVLESNIKVLPTWSTPV-----PCAPKPKCDKTHTC 411

Db 293 P-----SNTKVDTKVERKCCVECPAP-----317

QY 412 ELIGSPVFLPPKPKDMLTMRTPVTCVVDVSHEDPEVPMVVDVEMNATKTR 471

Db 318 --VAGSPVFLPPKPKDMLTMRTPVTCVVDVSHEDPEVPMVVDVEMNATKTR 375

QY 472 EQYNSTVAVSVLTVLHODMNGKYEKKVSNKALPAPIEKTISKAKQPREPQYTTIP 531

Db 376 EQFNSTFRVSVLTVLHODMNGKYEKKVSNKALPAPIEKTISKAKQPREPQYTTIP 435

QY 532 PSEDLTKQVSLTCLVKGFYPSDIAVENESNGQPPNNYKTPPVLDSDGSPFLYSLTV 591

Db 436 PSEDLTKQVSLTCLVKGFYPSDIAVENESNGQPPNNYKTPPVLDSDGSPFLYSLTV 495

QY 592 DKSROQGNVFCSCVWHEALHNHYTQKSLSPG 625

Db 496 DKSROQGNVFCSCVWHEALHNHYTQKSLSPG 529

RESULT 8  
PCT-US93-07422-4  
Sequence 4, Application PC/TUS9307422

GENERAL INFORMATION:

APPLICANT: Progenics Pharmaceuticals, Inc.

TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/07422

FILING DATE: 19930806

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/927,931

FILING DATE: 07-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 977-9809

TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 530 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ORIGINAL SOURCE: cDNA

ORGANISM: homo sapien

CELL TYPE: lymphocyte

PCT-US93-07422-4

Query Match 62.0%; Score 2116; DB 5; Length 530;

Best Local Similarity 68.9%; Pred. No. 7.3e-160; Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

```

Qy 1 MNRGVPFRHLILVQLALLPAATGKRVYLGKGDVIELTCTAOKKSIOFHMKNSQIK 60
Db 1 MNRGVPFRHLILVQLALLPAATGKRVYLGKGDVIELTCTAOKKSIOFHMKNSQIK 60
Qy 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
Qy 121 LVFGLTANSPTHLILOGSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Db 121 LVFGLTANSPTHLILOGSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Qy 181 TWCTCTVONQKKEFKIDIVLAFQKASSIYKKEGEQVESFPLAFTVEKLTGSGELMW 240
Db 181 TWCTCTVONQKKEFKIDIVLAFQKASSIYKKEGEQVESFPLAFTVEKLTGSGELMW 240
Qy 241 QAERASSKSMITFDLKNKEVSVRKVTQDPKLOMGKCLPLHLTLPOALPOYAG--SGNL 297
Db 241 QAERASSKSMITFDLKNKEVSVRKVTQDPKLOMGKCLPLHLTLPOALPOYAG--SGNL 297
Qy 298 TLAEAKTGKLEHENVLVVWRATOL-QKNLTCEVWGPTSPKXLMLSKLENKAKVSKREK 356
Db 298 TLAEAKTGKLEHENVLVVWRATOL-QKNLTCEVWGPTSPKXLMLSKLENKAKVSKREK 356
Qy 357 PWWALNPEAGMWOCILSDSGVLLSNIKVLPTWSTPV---PCPAPEKSCDKTHTCP 411
Db 357 PWWALNPEAGMWOCILSDSGVLLSNIKVLPTWSTPV---PCPAPEKSCDKTHTCP 411
Qy 412 ELGGSPVFLPPKPKOTLMTSRPEVTCVVVDVSHEDPEVKFMVYGVGVHNAKTKPR 471
Db 412 ELGGSPVFLPPKPKOTLMTSRPEVTCVVVDVSHEDPEVKFMVYGVGVHNAKTKPR 471
Qy 472 EEOYNSTYRVSVLTLVHODWLNKKEYKCKVSNKALPAPIEKTIISKAKGQPREPOVYTLR 531
Db 472 EEOYNSTYRVSVLTLVHODWLNKKEYKCKVSNKALPAPIEKTIISKAKGQPREPOVYTLR 531
Qy 532 PSRDELTKNOVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKITLV 591
Db 532 PSRDELTKNOVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKITLV 591
Qy 592 DKSRMOCNVFSCSVMEALHNHYTQKSLSLSPG 625
Db 592 DKSRMOCNVFSCSVMEALHNHYTQKSLSLSPG 625
Qy 496 DKSRMOCNVFSCSVMEALHNHYTQKSLSLSPG 529
Db 496 DKSRMOCNVFSCSVMEALHNHYTQKSLSLSPG 529

RESULT 9
US-08-477-460B-2
Sequence 2, Application US/08477460B
Patent No. 6034223
GENERAL INFORMATION:
APPLICANT: Progenics Pharmaceuticals, Inc.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED
NUMBER OF SEQUENCES: 9
TITLE OF INVENTION: CD4-GAMMA2 AND CD4-1G62 IMMUNOCONJUGATES, AND USES THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,460B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530

```

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/927,931

FILING DATE: 07-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/AJM

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 977-9805

TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-08-477-460B-2

## Query Match

Best Local Similarity 60.3%; Score 2059; DB 3; Length 432;

Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

Qy 1 MNRGVPFRHLILVQLALLPAATGKRVYLGKGDVIELTCTAOKKSIOFHMKNSQIK 60
Db 1 MNRGVPFRHLILVQLALLPAATGKRVYLGKGDVIELTCTAOKKSIOFHMKNSQIK 60
Qy 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
Db 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEYOL 120
Qy 121 LVFGLTANSPTHLILOGSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Db 121 LVFGLTANSPTHLILOGSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDQSG 180
Qy 181 TWCTCTVONQKKEFKIDIVLAFQKASSIYKKEGEQVESFPLAFTVEKLTGSGELMW 240
Db 181 TWCTCTVONQKKEFKIDIVLAFQKASSIYKKEGEQVESFPLAFTVEKLTGSGELMW 240
Qy 241 QAERASSKSMITFDLKNKEVSVRKVTQDPKLOMGKCLPLHLTLPOALPOYAGSGNLTLA 300
Db 241 QAERASSKSMITFDLKNKEVSVRKVTQDPKLOMGKCLPLHLTLPOALPOYAGSGNLTLA 300
Qy 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEVWGPTSPKXLMLSKLENKAKVSKREKPMV 360
Db 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEVWGPTSPKXLMLSKLENKAKVSKREKPMV 360
Qy 361 LNPAGMWOCILSDSGVLLSNIKVLPTWSTPVPCPAPEKSCDKTHTCPDELIGSPVVF 420
Db 361 LNPAGMWOCILSDSGVLLSNIKVLPTWSTPVPCPAPEKSCDKTHTCPDELIGSPVVF 420
Qy 421 LFPKPKOTLMTSRPEVTCVVVDVSHEDPEVKFMVYGVGVHNAKTKPREDFNSYF 226
Db 421 LFPKPKOTLMTSRPEVTCVVVDVSHEDPEVKFMVYGVGVHNAKTKPREDFNSYF 226
Qy 481 VVSVLTIVHODWLNKKEYKCKVSNKALPAPIEKTIISKAKGQPREPOVYTLRPSRDELTKN 540
Db 481 VVSVLTIVHODWLNKKEYKCKVSNKALPAPIEKTIISKAKGQPREPOVYTLRPSRDELTKN 540
Qy 541 QVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKITLVDSRWQGN 600
Db 541 QVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKITLVDSRWQGN 600
Qy 601 VFCSSVMEALHNHYTQKSLSLSPG 625
Db 601 VFCSSVMEALHNHYTQKSLSLSPG 625
Qy 407 VFCSSVMEALHNHYTQKSLSLSPG 431
Db 407 VFCSSVMEALHNHYTQKSLSLSPG 431

RESULT 10

```

```
US-08-379-516-2
; Sequence 2, Application US/08379516
; Patent No. 6083478
; GENERAL INFORMATION:
; APPLICANT: Allaway, Graham P.
; APPLICANT: Maddon, Paul J.
; TITLE OF INVENTION: No. 6083478-Peptidyl Moiety-Conjugated CD4-Gamma2 and CD4-IgG2
; TITLE OF INVENTION: Immunconjugates and Uses Thereof
; FILE REFERENCE: 41215-A-PCT-US
; CURRENT APPLICATION NUMBER: US/08/379,516
; CURRENT FILING DATE: 1996-06-10
; EARLIER APPLICATION NUMBER: PCT/US93/07422
; EARLIER FILING DATE: 1993-08-06
; EARLIER APPLICATION NUMBER: 07/927,931
; EARLIER FILING DATE: 1992-08-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 432
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-379-516-2

Query Match      60.3%; Score 2059; DB 3; Length 432;
Best Local Similarity 65.6%; Pred. No. 1,8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVPRHLLVLQALPAATQGNKVLGKGGDVELTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVLQALPAATQGNKVLGKGGDVELTCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLITKNKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLITKNKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
QY 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
QY 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
QY 241 QAERASSKSWITFDLKNKREVSVRVYQDPKLOMGKRLPLHLTLPLQALPQYAGSGNLTLLA 300
DB 241 QAERASSKSWITFDLKNKREVSVRVYQDPKLOMGKRLPLHLTLPLQALPQYAGSGNLTLLA 300
QY 211 ----- 210
DB 211 ----- 210
QY 301 LEAKTGKLBQEVNLVVRATQLOKNLTCFVWGPTSPLYMLSLKLENKAKVSKREKPVWV 360
DB 301 LEAKTGKLBQEVNLVVRATQLOKNLTCFVWGPTSPLYMLSLKLENKAKVSKREKPVWV 360
QY 211 ----- 210
DB 211 ----- 210
QY 361 LNPBAGMOCLLSDSGVLLBSNLIKVLPTWSTVPYPCAPBPBKSCDKTHTCPELLGSPSVF 420
DB 361 LNPBAGMOCLLSDSGVLLBSNLIKVLPTWSTVPYPCAPBPBKSCDKTHTCPELLGSPSVF 420
QY 211 -----EC-----PPCPAPP-----VAGSPVF 226
DB 211 -----EC-----PPCPAPP-----VAGSPVF 226
QY 421 LFPKPKDQTLMIISTPBEVTCVVVDVSHEDPEVKFKNWYVDGVEVHNATKREEQYNSTYR 480
DB 421 LFPKPKDQTLMIISTPBEVTCVVVDVSHEDPEVKFKNWYVDGVEVHNATKREEQYNSTYR 480
QY 227 LFPKPKDQTLMIISTPBEVTCVVVDVSHEDPEVKFKNWYVDGVEVHNATKREEQYNSTYR 286
DB 227 LFPKPKDQTLMIISTPBEVTCVVVDVSHEDPEVKFKNWYVDGVEVHNATKREEQYNSTYR 286
QY 481 VVSVLTVLHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 540
DB 481 VVSVLTVLHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 540
QY 287 VVSVLTVLHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 346
DB 287 VVSVLTVLHODMLNGEKYCKCKVSNKALPAPLEKTIISKAKGQPREPQVYTLPPSRDELTKN 346
QY 541 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSGSPFLYSKLTVDKSRMQGN 600
DB 541 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSGSPFLYSKLTVDKSRMQGN 600
QY 347 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSGSPFLYSKLTVDKSRMQGN 406
DB 347 QVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPRVLDSGSPFLYSKLTVDKSRMQGN 406
QY 601 VFGSVMHEALAHNHYTQKSLSLSPG 625
DB 601 VFGSVMHEALAHNHYTQKSLSLSPG 625
QY 407 VFGSVMHEALAHNHYTQKSLSLSPG 431
DB 407 VFGSVMHEALAHNHYTQKSLSLSPG 431
```

```
US-09-329-916-2
; Sequence 2, Application US/09329916
; Patent No. 6177549
; GENERAL INFORMATION:
; APPLICANT: Progenics Pharmaceuticals, Inc.
; APPLICANT: NON-PEPTIDYL MOIETY-CONJUGATED
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IgG2 IMMUNOCONJUGATES, AND USES THEREOF
; TITLE OF INVENTION: CD4-GAMMA2 AND CD4-IgG2 IMMUNOCONJUGATES, AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/329,916
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/477,460
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 07/927,931
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPM/AJM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 977-9809
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; CELL TYPE: lymphocyte
US-09-329-916-2

Query Match      60.3%; Score 2059; DB 3; Length 432;
Best Local Similarity 65.6%; Pred. No. 1,8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVPRHLLVLQALPAATQGNKVLGKGGDVELTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVLQALPAATQGNKVLGKGGDVELTCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLITKNKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLITKNKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
QY 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
QY 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWCTCTVLQGNQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240
QY 241 QAERASSKSWITFDLKNKREVSVRVYQDPKLOMGKRLPLHLTLPLQALPQYAGSGNLTLLA 300
DB 241 QAERASSKSWITFDLKNKREVSVRVYQDPKLOMGKRLPLHLTLPLQALPQYAGSGNLTLLA 300
QY 211 ----- 210
DB 211 ----- 210
```

Qy 301 LEAKTKLHGEVNLVWRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360  
Db 211 ----- 210  
Qy 361 LNPEAGMOCCLSDSGVLLSNIKVLPTWSTPVPCEPAPEPKSCDKHTCPCLLGGPSVF 420  
Db 211 -----EC-----PPCPAP-----VAGPSVF 226  
Qy 421 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVNNAKTKPREBOYNSIYR 480  
Db 227 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVNNAKTKPREBOYNSIYR 286  
Qy 481 VVSULTLVHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN 540  
Db 287 VVSULTLVHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN 346  
Qy 541 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRMOCN 600  
Db 347 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRMOCN 406  
Qy 601 VFSCSVNHEALHNHYTOKSLSLSPG 625  
Db 407 VFSCSVNHEALHNHYTOKSLSLSPG 431

## RESULT 12

US-08-485-372A-2

Sequence 2, Application US/08485372A

Patent No. 6187748

GENERAL INFORMATION:

APPLICANT: Beaudry, Gary A.

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 37690-11-A

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-08-485-372A-2

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,227

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 37690-11-A

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 432 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: homo sapien

CELL TYPE: lymphocyte

US-08-485-372A-2

Query Match 60.3%; Score 2059; DB 3; Length 432;  
Best Local Similarity 65.6%; Pred. No. 1,8e-155;  
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

Qy 1 MNRGVPFRHLILVQLALPAPATQGNKVLGKGDVLTCTASQKKSIOFHMNSQIK 60  
Db 1 MNRGVPFRHLILVQLALPAPATQGNKVLGKGDVLTCTASQKKSIOFHMNSQIK 60  
Qy 61 ILGNQGSFLTKGPEKNDNRDSDRSRSLMDQGNFPLIKNLKEDSDTYICEVEDQKEVQL 120  
Db 61 ILGNQGSFLTKGPEKNDNRDSDRSRSLMDQGNFPLIKNLKEDSDTYICEVEDQKEVQL 120  
Qy 121 LVFGLTANSPTHLILQGSLLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLELDSDG 180  
Db 121 LVFGLTANSPTHLILQGSLLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLELDSDG 180  
Qy 181 TWCTCTVLQNGKVFVKIDIVLAFQKASSIVYKKEGQVEFSPFLATVEXLTGSGELMW 240  
Db 181 TWCTCTVLQNGKVFVKIDIVLAFQKASSIVYKKEGQVEFSPFLATVEXLTGSGELMW 240  
Qy 241 QABRASSKSWITFDLKNKEVSVRVQDRKLGKTLPLHLTLPQALPOYAGSGLTLA 300  
Db 211 ----- 210  
Qy 301 LEAKTKLHGEVNLVWRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360  
Db 211 ----- 210  
Qy 361 LNPEAGMOCCLSDSGVLLSNIKVLPTWSTPVPCEPAPEPKSCDKHTCPCLLGGPSVF 420  
Db 211 -----EC-----PPCPAP-----VAGPSVF 226  
Qy 421 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVNNAKTKPREBOYNSIYR 480  
Db 227 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVNNAKTKPREBOYNSIYR 286  
Qy 481 VVSULTLVHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN 540  
Db 287 VVSULTLVHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN 346  
Qy 541 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRMOCN 600  
Db 347 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGFFLYSKLTVDKSRMOCN 406  
Qy 601 VFSCSVNHEALHNHYTOKSLSLSPG 625  
Db 407 VFSCSVNHEALHNHYTOKSLSLSPG 431

## RESULT 13

US-09-409-006A-2

Sequence 2, Application US/09409006A

Patent No. 6342586

GENERAL INFORMATION:

APPLICANT: Progenics Pharmaceuticals, Inc.

TITLE OF INVENTION: NON-PEPTIDYL NOIETY-CONJUGATED

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cooper &amp; Dunham

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/409,006A

FILING DATE: 29-SEP-1999

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,931
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41215-A-PCT/JPW/ALM
TELEPHONE: (212) 977-9550
TELEFAX: (212) 977-9809
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-09-409-006A-2

```

```

Query Match      60.3%; Score 2059; DB 4; Length 432;
Best Local Similarity 65.6%; Pred. No. 1.8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

```

QY 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASQKKSIOFHKNSNOIK 60
DB 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASQKKSIOFHKNSNOIK 60
QY 61 ILNGQSFLLTKGSPSKLNDRAISRSLMDQGNPFLIINKLIEBDDTYICEVEDQKEVOL 120
DB 61 ILNGQSFLLTKGSPSKLNDRAISRSLMDQGNPFLIINKLIEBDDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVOCSPRGKNIQGGKTLVSQLELDGSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVOCSPRGKNIQGGKTLVSQLELDGSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVFSPLAFTVEKLTGSGELMW 240
QY 241 QABRASSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QABRASSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLEHVNVLVVRATQLOKNLTCVWGPTSPKMLSLKLENKAQVSKREKPVWV 360
DB 301 LEAKTGKLEHVNVLVVRATQLOKNLTCVWGPTSPKMLSLKLENKAQVSKREKPVWV 360
QY 421 LFPKPKDITLMISTPEVTCVVDVSHEDPEYKKNVVDGVEYHNAKTKRREQYNSTYR 480
DB 421 LFPKPKDITLMISTPEVTCVVDVSHEDPEYKKNVVDGVEYHNAKTKRREQYNSTYR 480
QY 481 VWSVLTAVLHODMLNGEKYCKVSNKALPAPIEKTISAKGQPREPOVYTLPPSRDELTKN 540
DB 481 VWSVLTAVLHODMLNGEKYCKVSNKALPAPIEKTISAKGQPREPOVYTLPPSRDELTKN 540
QY 541 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMOQN 600
DB 541 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMOQN 600
QY 601 VFSCSVHHEALHNHYTQKSLSLSPG 625
DB 601 VFSCSVHHEALHNHYTQKSLSLSPG 625
QY 407 VFSCSVHHEALHNHYTQKSLSLSPG 431
DB 407 VFSCSVHHEALHNHYTQKSLSLSPG 431

```

```

RESULT 14
US-08-484-681-2
Sequence 2, Application US/08484681
Patent No. 6451313
GENERAL INFORMATION:
APPLICANT: Beaudry, Gary A.
APPLICANT: Maddon, Paul J.
TITLE OF INVENTION: CD4-GAMMA2 CD4-IGG2 CHIMERAS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 37690-II-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: homo sapien
CELL TYPE: lymphocyte
US-08-484-681-2

```

```

Query Match      60.3%; Score 2059; DB 4; Length 432;
Best Local Similarity 65.6%; Pred. No. 1.8e-155;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

```

QY 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASQKKSIOFHKNSNOIK 60
DB 1 MNRGVPRHLLVLTQALPAATQGNKVVLGKKGDVLTCTASQKKSIOFHKNSNOIK 60
QY 61 ILNGQSFLLTKGSPSKLNDRAISRSLMDQGNPFLIINKLIEBDDTYICEVEDQKEVOL 120
DB 61 ILNGQSFLLTKGSPSKLNDRAISRSLMDQGNPFLIINKLIEBDDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVOCSPRGKNIQGGKTLVSQLELDGSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVOCSPRGKNIQGGKTLVSQLELDGSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGEQVFSPLAFTVEKLTGSGELMW 240
QY 241 QABRASSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QABRASSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLEHVNVLVVRATQLOKNLTCVWGPTSPKMLSLKLENKAQVSKREKPVWV 360
DB 301 LEAKTGKLEHVNVLVVRATQLOKNLTCVWGPTSPKMLSLKLENKAQVSKREKPVWV 360
QY 421 LFPKPKDITLMISTPEVTCVVDVSHEDPEYKKNVVDGVEYHNAKTKRREQYNSTYR 480
DB 421 LFPKPKDITLMISTPEVTCVVDVSHEDPEYKKNVVDGVEYHNAKTKRREQYNSTYR 480
QY 481 VWSVLTAVLHODMLNGEKYCKVSNKALPAPIEKTISAKGQPREPOVYTLPPSRDELTKN 540
DB 481 VWSVLTAVLHODMLNGEKYCKVSNKALPAPIEKTISAKGQPREPOVYTLPPSRDELTKN 540
QY 541 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMOQN 600
DB 541 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMOQN 600
QY 601 VFSCSVHHEALHNHYTQKSLSLSPG 625
DB 601 VFSCSVHHEALHNHYTQKSLSLSPG 625
QY 407 VFSCSVHHEALHNHYTQKSLSLSPG 431
DB 407 VFSCSVHHEALHNHYTQKSLSLSPG 431

```

```

Qy      361  LNPBAGMIGCOLLSDSGVLLSNIKVLPWSPFVPCPAPEPKSCDKHTHCPELLGSPVF 420
      : ||| ||| ||| |||
Db      211  -----EC-----PCCPAP-----VAGSVF 226
      : ||| ||| ||| |||
Qy      421  LPPPKKDTLMSRTEPEVTCVVVDVSHEDPEVKFMYVGYGVNNAKTKPREQVNSTYR 480
      : ||| ||| ||| |||
Db      227  LPPPKKDTLMSRTEPEVTCVVVDVSHEDPEVQFMWYVDGYGVNNAKTKPREQFNSTFR 286
      : ||| ||| ||| |||
Qy      481  VWSVLTVLHDDMGNGEYKCAVSNKLPAPRIEKTSKAGQPREPOVYTLPPRSDELTKN 540
      : ||| ||| ||| |||
Db      287  VWSVLTVVMODMNGEYKCAVSNKLPAPRIEKTSKAGQPREPOVYTLPPRSDEMTKN 346
      : ||| ||| ||| |||
Qy      541  QVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPRLVLDSDSFLYSKLTVDKSRMOQN 600
      : ||| ||| ||| |||
Db      347  QVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPRLVLDSDSFLYSKLTVDKSRMOQN 406
      : ||| ||| ||| |||
Qy      601  VFGSCVNHGALHNHYTKSLSPG 625
      : ||| ||| ||| |||
Db      407  VFGSCVNHGALHNHYTKSLSPG 431

```

```

1  MNRGVPFPHLLLVLDLALPLPAATGKVKVLLGKKGTVELTCTASQKKSQIFHWKNSNQIK 60
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032

```

```

Db 1 MNRGVPRHLLVLOALPRAATQGGKVVLGKKGDVLTCTASQKKSIOFMKMSNOJK 60
Qy 61 ILGNQGSFLTGGPSKLNDRADSRRSIMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVL 12
Db 61 ILGNQGSFLTGGPSKLNDRADSRRSIMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVL 12
Qy 121 LVFGLTANSPTTHLLOGOSLTLTLESPQGSAPSVQCSPRGXNIQGGKTLVSQLELQDSDG 18
Db 121 LVFGLTANSPTTHLLOGOSLTLTLESPQGSAPSVQCSPRGXNIQGGKTLVSQLELQDSDG 18
Qy 181 TWTTCTVLQONKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFLAFTVEKLTGSGELTW 24
Db 181 TWTTCTVLQONKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFLAFTVEKLTGSGELTW 24
Qy 241 QABEASSKSMITFDLKNKEVSKRATQDPKQLQMGKCLP LHLTLPQALPQVAGSGLTLA 30
Db 241 QABEASSKSMITFDLKNKEVSKRATQDPKQLQMGKCLP LHLTLPQALPQVAGSGLTLA 30
Qy 301 LEAKTGKLHQBENLVYMRATOLQKNLTCEVWGFTSPKMLSLKENKAQVSKREKPVWV 36
Db 301 LEAKTGKLHQBENLVYMRATOLQKNLTCEVWGFTSPKMLSLKENKAQVSKREKPVWV 36
Qy 361 LNPBAGMWOCLLDSQGVLLSEINIKVLPMTWSTPVRCPAPBPKSCDKHTCPBELLGSPV 42
Db 361 LNPBAGMWOCLLDSQGVLLSEINIKVLPMTWSTPVRCPAPBPKSCDKHTCPBELLGSPV 42
Qy 421 LFPKPCKDTLMSRTPEVTCVVVDVSHEDPEVKFNNYVYDGEVHNAAKTRPEEOYNSTVR 48
Db 421 LFPKPCKDTLMSRTPEVTCVVVDVSHEDPEVKFNNYVYDGEVHNAAKTRPEEOYNSTVR 48
Qy 481 LFPKPCKDTLMSRTPEVTCVVVDVSHEDPEVKFNNYVYDGEVHNAAKTRPEEOYNSTVR 48
Db 481 LFPKPCKDTLMSRTPEVTCVVVDVSHEDPEVKFNNYVYDGEVHNAAKTRPEEOYNSTVR 48
Qy 541 QVSLTCLVUKFYPSDIAVEMESNGCPENNYKTTPRVLDSDGFFLYSKLTVDKSRWQGN 60
Db 541 QVSLTCLVUKFYPSDIAVEMESNGCPENNYKTTPRVLDSDGFFLYSKLTVDKSRWQGN 60
Qy 601 VFGSVMHEALHNHNTQKSLSPG 625
Db 601 VFGSVMHEALHNHNTQKSLSPG 625
Qy 625 VFGSVMHEALHNHNTQKSLSPG 625
Db 625 VFGSVMHEALHNHNTQKSLSPG 625

```

RESULT 16  
 US-08-417-495-6  
 Sequence 6, Application US/08417495  
 Patent No. 5845728  
 GENERAL INFORMATION:  
 APPLICANT: Seed, Brian et al.  
 TITLE OF INVENTION: Redirection of Cellular Immunity by  
 TITLE OF INVENTION: Receptor  
 NUMBER OF SEQUENCES: 27  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson  
 STREET: 225 Franklin Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110-2804  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM PS/2 Model 50 or 55SX  
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
 SOFTWARE: WordPerfect (Version 5.0)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/417,495  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION: 435  
 APPLICATION NUMBER: US/08/203,866  
 FILING DATE:  
 APPLICATION NUMBER: US/07/847,566

FILING DATE: 07/665,961  
APPLICATION NUMBER: 07/665,961  
FILING DATE: Match 7, 1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/119002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-417-495-6

Query Match 59.7%; Score 2036.5; DB 2; Length 532;  
Best Local Similarity 98.3%; Pred. No. 1.5e-153;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPRHLLVLTALPAPATQGNKVLGKKGDTVELTCTASOKKSIOFHMKNSNQIK 60  
DB 1 MNRGVPRHLLVLTALPAPATQGNKVLGKKGDTVELTCTASOKKSIOFHMKNSNQIK 60  
QY 61 ILGNQGSFLLTKGPSKLNDRADRSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEVL 120  
DB 61 ILGNQGSFLLTKGPSKLNDRADRSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEVL 120  
QY 121 LVFGLTRANSDTHLLOQSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTRANSDTHLLOQSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTCTVLQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240  
DB 181 TWCTCTVLQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240  
QY 241 QAEARASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
DB 241 QAEARASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
QY 301 LEAKTGKLEHENVLVVWRATQLOKNTCEVWGPTSPKMLSLKLENKEAKVSKREKPVW 360  
DB 301 LEAKTGKLEHENVLVVWRATQLOKNTCEVWGPTSPKMLSLKLENKEAKVSKREKPVW 360  
QY 361 LNPBAGMOCCLSDSGQVLLSNIKVLPTWSTPVPCEAPPEKSC 404  
DB 361 LNPBAGMOCCLSDSGQVLLSNIKVLPTWSTPVPCEAPPEKSC 404

RESULT 17  
US-08-284-391B-6  
Sequence 6, Application US/08284391B  
Patent No. 5851828  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanue, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-6

Query Match 59.7%; Score 2036.5; DB 2; Length 532;  
Best Local Similarity 98.3%; Pred. No. 1.5e-153;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPRHLLVLTALPAPATQGNKVLGKKGDTVELTCTASOKKSIOFHMKNSNQIK 60  
DB 1 MNRGVPRHLLVLTALPAPATQGNKVLGKKGDTVELTCTASOKKSIOFHMKNSNQIK 60  
QY 61 ILGNQGSFLLTKGPSKLNDRADRSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEVL 120  
DB 61 ILGNQGSFLLTKGPSKLNDRADRSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEEVL 120  
QY 121 LVFGLTRANSDTHLLOQSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTRANSDTHLLOQSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTCTVLQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240  
DB 181 TWCTCTVLQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240  
QY 241 QAEARASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
DB 241 QAEARASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
QY 301 LEAKTGKLEHENVLVVWRATQLOKNTCEVWGPTSPKMLSLKLENKEAKVSKREKPVW 360  
DB 301 LEAKTGKLEHENVLVVWRATQLOKNTCEVWGPTSPKMLSLKLENKEAKVSKREKPVW 360  
QY 361 LNPBAGMOCCLSDSGQVLLSNIKVLPTWSTPVPCEAPPEKSC 404  
DB 361 LNPBAGMOCCLSDSGQVLLSNIKVLPTWSTPVPCEAPPEKSC 404

RESULT 18  
US-09-218-950-6  
Sequence 6, Application US/09218950  
Patent No. 6284240  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles



```
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-6

Query Match 59.7%; Score 2036.5; DB 3; Length 532;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLLVQLALLPAATGKNKYLGGKGTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLLVQLALLPAATGKNKYLGGKGTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLXIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLXIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
QY 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVRKYVQDPKLGKGLPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QAERASSSKSWITFDLKNKEVSVRKYVQDPKLGKGLPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLGHEVNLVVMRATOLQKRLTCEVWGPTSPKMLSLKLENKAKYSKREKPYMV 360
DB 301 LEAKTGKLGHEVNLVVMRATOLQKRLTCEVWGPTSPKMLSLKLENKAKYSKREKPYMV 360
```

```
RESULT 19
PCT-US92-01785-6
Sequence 6, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Mordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-01785-6

Query Match 59.7%; Score 2036.5; DB 5; Length 532;
Best Local Similarity 98.3%; Pred. No. 1.5e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLLVQLALLPAATGKNKYLGGKGTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLLVQLALLPAATGKNKYLGGKGTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLXIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLXIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSSTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
QY 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWTCIVLONOKKVEKIDIVLAFQKASIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVRKYVQDPKLGKGLPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QAERASSSKSWITFDLKNKEVSVRKYVQDPKLGKGLPLHLTLPOALPOYAGSGNLTIA 300
```

QY 301 LEAKTGKHOEVNLYVMRATOLQKNTCEVWGPTSPKMLSLKLEKKAQVSKREKPVWV 360  
| | | | |  
DB 301 LEAKTGKHOEVNLYVMRATOLQKNTCEVWGPTSPKMLSLKLEKKAQVSKREKPVWV 360  
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPKSC 404  
| | | | |  
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVAH---DPKLC 401

RESULT 20  
PCT-US95-00454-6  
; Sequence 6, Application PC/TUS9500454  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian et al.  
; TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected  
; TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/00454  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: March 6, 1992  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/247001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 532 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-00454-6

Query Match 59.7%; Score 2036.5; DB 5; Length 532;  
Best Local Similarity 98.3%; Pred. No. 1.5e-153;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLILVQLALLPAATGKNVVLGKKGDTVELTCTASOKKSIQFHWKNSNQIK 60  
| | | | |  
DB 1 MNRGVPFRHLILVQLALLPAATGKNVVLGKKGDTVELTCTASOKKSIQFHWKNSNQIK 60  
QY 61 ILNGQSFILTKGSPKLNDRADSRSLMDQGNFPIILNKLEISDPTICVEEDQKEEVOL 120  
| | | | |  
DB 61 ILNGQSFILTKGSPKLNDRADSRSLMDQGNFPIILNKLEISDPTICVEEDQKEEVOL 120  
QY 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVSOLELQDSG 180  
| | | | |  
DB 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVSOLELQDSG 180

QY 181 TWTCVTLQNOQKVEPFIDIVLAFQKASIVYKKEGOVEFSFPLAFVTEKLTGSGELMW 240  
| | | | |  
DB 181 TWTCVTLQNOQKVEPFIDIVLAFQKASIVYKKEGOVEFSFPLAFVTEKLTGSGELMW 240  
QY 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLQMGKPLHLTLPOLPOYAGSGNLTJA 300  
| | | | |  
DB 241 QAERASSSKSWITFDLKNKEVSVKRYTODPKLQMGKPLHLTLPOLPOYAGSGNLTJA 300  
QY 301 LEAKTGKHOEVNLYVMRATOLQKNTCEVWGPTSPKMLSLKLEKKAQVSKREKPVWV 360  
| | | | |  
DB 301 LEAKTGKHOEVNLYVMRATOLQKNTCEVWGPTSPKMLSLKLEKKAQVSKREKPVWV 360  
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPKSC 404  
| | | | |  
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVAH---DPKLC 401

RESULT 21  
US-08-417-495-4  
; Sequence 4, Application US/08417495  
; Patent No. 5843728  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian et al.  
; TITLE OF INVENTION: Redirection of Cellular Immunity by  
; TITLE OF INVENTION: Receptor  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/417,495  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/203,866  
; FILING DATE:  
; APPLICATION NUMBER: US/07/847,566  
; FILING DATE:  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/119002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-417-495-4

Query Match 59.7%; Score 2036.5; DB 2; Length 575;  
Best Local Similarity 98.3%; Pred. No. 1.7e-153;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLILVQLALLPAATGKNVVLGKKGDTVELTCTASOKKSIQFHWKNSNQIK 60  
| | | | |

```

Db      1  MNRGVPRHLLVLVQLALLPAATQGNKRVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY      61  ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINIKLKIEDSDTYICEVEDQKEEVQL 120
        61  ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINIKLKIEDSDTYICEVEDQKEEVQL 120
Db      121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
        121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
QY      121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
        121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
Db      181  TWCTVLOQNKKEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
        181  TWCTVLOQNKKEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
QY      241  QAERASSSKSWITFDLNKKEVSVKRVTDPLQNGKKLPLHLTLPOLPOYAGSGNLTILA 300
        241  QAERASSSKSWITFDLNKKEVSVKRVTDPLQNGKKLPLHLTLPOLPOYAGSGNLTILA 300
Db      301  LEATGKLHGEVNLVWMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKYSKEKEPVWV 360
        301  LEATGKLHGEVNLVWMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKYSKEKEPVWV 360
QY      361  LNPAGMWQCLLSDSGVLLSESNIKVLPMTSTPVPCPAPEKSC 404
        361  LNPAGMWQCLLSDSGVLLSESNIKVLPMTSTPVPHA--DBKLC 401
Db      361  LNPAGMWQCLLSDSGVLLSESNIKVLPMTSTPVPHA--DBKLC 401

RESULT 22
US-08-284-391B-4
; Sequence 4, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-284-391B-4

Query Match      59.7%; Score 2036.5; DB 2; Length 575;
Best Local Similarity 98.3%; Pred. No. 1.7e-153;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Db      1  MNRGVPRHLLVLVQLALLPAATQGNKRVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60
QY      61  ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINIKLKIEDSDTYICEVEDQKEEVQL 120
        61  ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINIKLKIEDSDTYICEVEDQKEEVQL 120
Db      121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
        121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
QY      121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
        121  LVFGLTANSDTHLLQGSGLTTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
Db      181  TWCTVLOQNKKEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
        181  TWCTVLOQNKKEFKIDIVVLAFOKASSIYKKEGQVEFSPLAFTVEKLTGSGELMW 240
QY      241  QAERASSSKSWITFDLNKKEVSVKRVTDPLQNGKKLPLHLTLPOLPOYAGSGNLTILA 300
        241  QAERASSSKSWITFDLNKKEVSVKRVTDPLQNGKKLPLHLTLPOLPOYAGSGNLTILA 300
Db      241  QAERASSSKSWITFDLNKKEVSVKRVTDPLQNGKKLPLHLTLPOLPOYAGSGNLTILA 300
        241  QAERASSSKSWITFDLNKKEVSVKRVTDPLQNGKKLPLHLTLPOLPOYAGSGNLTILA 300
QY      301  LEATGKLHGEVNLVWMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKYSKEKEPVWV 360
        301  LEATGKLHGEVNLVWMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKYSKEKEPVWV 360
Db      301  LEATGKLHGEVNLVWMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKYSKEKEPVWV 360
        301  LEATGKLHGEVNLVWMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKYSKEKEPVWV 360
QY      361  LNPAGMWQCLLSDSGVLLSESNIKVLPMTSTPVPCPAPEKSC 404
        361  LNPAGMWQCLLSDSGVLLSESNIKVLPMTSTPVPHA--DBKLC 401
Db      361  LNPAGMWQCLLSDSGVLLSESNIKVLPMTSTPVPHA--DBKLC 401

RESULT 23
US-09-218-950-4
; Sequence 4, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566

```

FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elding, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 575 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-4

Query Match 59.7%; Score 2036.5; DB 3; Length 575;  
Best Local Similarity 98.3%; Pred. No. 1.7e-153;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLTIEDSDTYICVEBQKEVOL 120  
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLTIEDSDTYICVEBQKEVOL 120  
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLEIDSG 180  
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLEIDSG 180  
QY 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240  
DB 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240  
QY 241 QAEARASSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPOLPOYAGSGNLTIA 300  
DB 241 QAEARASSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPOLPOYAGSGNLTIA 300  
QY 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360  
DB 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360  
QY 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTVPVPCAPPEKSC 404  
DB 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTVPVPCAPPEKSC 404

RESULT 24  
PCT-US92-01785-4  
Sequence 4, Application PC/TUS9201785  
GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation  
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor  
TITLE OF INVENTION: Chimerae  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/01785  
FILING DATE: 19920306  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/119002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 575 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-01785-4

Query Match 59.7%; Score 2036.5; DB 5; Length 575;  
Best Local Similarity 98.3%; Pred. No. 1.7e-153;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLTIEDSDTYICVEBQKEVOL 120  
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLTIEDSDTYICVEBQKEVOL 120  
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLEIDSG 180  
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLEIDSG 180  
QY 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240  
DB 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFVTEKLTGSGELMW 240  
QY 241 QAEARASSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPOLPOYAGSGNLTIA 300  
DB 241 QAEARASSKSWITFDLKNKEVSVKRYTODPKLOMGKKLPLHLTLPOLPOYAGSGNLTIA 300  
QY 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360  
DB 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360  
QY 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTVPVPCAPPEKSC 404  
DB 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTVPVPCAPPEKSC 404

RESULT 25  
PCT-US95-00454-4  
Sequence 4, Application PC/TUS9500454  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected  
TITLE OF INVENTION: Cells by Chimeraic CD4 Receptor-  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/00454  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/847,566  
FILING DATE: March 6, 1992  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 575 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-00454-4

Query Match 59.7%; Score 2036.5; DB 5; Length 575;  
Best Local Similarity 98.3%; Pred. No. 1,76-153;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;  
QY 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVETLCTASQKSIQPHMNSNQIK 60  
DB 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVETLCTASQKSIQPHMNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120  
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
QY 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLOQKKEVFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240  
DB 181 TWCTVLOQKKEVFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240  
QY 241 QAEBAASSKSWITFDLKNKEVSVKRVTDPRLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
DB 241 QAEBAASSKSWITFDLKNKEVSVKRVTDPRLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
QY 301 LEATGTGLHGEVNLVVMRATOLQKNLTCEWGPSPKLMSTLKENKEAVSKREKPVWV 360  
DB 301 LEATGTGLHGEVNLVVMRATOLQKNLTCEWGPSPKLMSTLKENKEAVSKREKPVWV 360  
QY 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404  
DB 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404

RESULT 26  
US-08-417-495-5  
Sequence 5, Application US/08417495  
Patent No. 5843728  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
TITLE OF INVENTION: Receptor Chimera

NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,495  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/203,866  
FILING DATE:  
APPLICATION NUMBER: US/07/847,566  
FILING DATE:  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/119002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-417-495-5

Query Match 59.5%; Score 2032.5; DB 2; Length 462;  
Best Local Similarity 98.0%; Pred. No. 2,66-153;  
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;  
QY 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVETLCTASQKSIQPHMNSNQIK 60  
DB 1 MNRGVPRHLLVQLALLPAATQGNKVLGKGDVETLCTASQKSIQPHMNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120  
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
QY 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSPTHLLOQGSITLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLOQKKEVFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240  
DB 181 TWCTVLOQKKEVFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240  
QY 241 QAEBAASSKSWITFDLKNKEVSVKRVTDPRLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
DB 241 QAEBAASSKSWITFDLKNKEVSVKRVTDPRLOMGKKLPLHLTLPQALPOYAGSGNLTIA 300  
QY 301 LEATGTGLHGEVNLVVMRATOLQKNLTCEWGPSPKLMSTLKENKEAVSKREKPVWV 360  
DB 301 LEATGTGLHGEVNLVVMRATOLQKNLTCEWGPSPKLMSTLKENKEAVSKREKPVWV 360  
QY 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404  
DB 361 LNPAGMOCCLSDSGOVLLESNIKVLPTWSTVPAPAPKSC 404

```
RESULT 27
US-08-284-391B-5
; Sequence 5, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-284-391B-5

Query Match      59.5%; Score 2032.5; DB 2; Length 462;
Best Local Similarity 98.0%; Pred. No. 2.6e-153;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
```

```

QY 241 QAEKASSKSWITFDLKNKEVSVKRVTDPPKLOMGKLLPLHLTLPOALPOYAGSGNLTLLA 300
DB 241 QAEKASSKSWITFDLKNKEVSVKRVTDPPKLOMGKLLPLHLTLPOALPOYAGSGNLTLLA 300
QY 301 LEAKTGKLEHVEVNLVVMRATOLQKULTCFVWGPTSPKMLSLKLENKAKVSRKRPVWY 360
DB 301 LEAKTGKLEHVEVNLVVMRATOLQKULTCFVWGPTSPKMLSLKLENKAKVSRKRPVWY 360
QY 361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTPVPCCPAPPEKSC 404
DB 361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTPVPVHA---DPQLC 401

RESULT 28
US-09-218-950-5
; Sequence 5, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-218-950-5

Query Match      59.5%; Score 2032.5; DB 3; Length 462;
Best Local Similarity 98.0%; Pred. No. 2.6e-153;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
```

Db 1 MNRGVPFRHLLLVLTQALLPAATQGNKVVLGKGDVTELCTASQKKSIOFHMKNNOIK 60  
Qy 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
Qy 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVOCRSRPGKNIOGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVOCRSRPGKNIOGKTLVSQLELDQSG 180  
Qy 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEQVEFSPFLATFVEKLTGSGELMW 240  
Db 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEQVEFSPFLATFVEKLTGSGELMW 240  
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPOALPOYAGSGNLTLA 300  
Db 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPOALPOYAGSGNLTLA 300  
Qy 301 LEATGKLHQBVLVVMRATOLQKNTCEVWGPTSPKLMSTKLENKAKVSKREKPYWV 360  
Db 301 LEATGKLHQBVLVVMRATOLQKNTCEVWGPTSPKLMSTKLENKAKVSKREKPYWV 360  
Qy 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVAH--DPQLC 401  
Db 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVAH--DPQLC 401

RESULT 29  
PCT-US92-01785-5  
Sequence 5, Application PC/TUS9201785

GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation  
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/01785  
FILING DATE: 19920306  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/119002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-01785-5

Query Match 59.5%; Score 2032.5; DB 5; Length 462;  
Best Local Similarity 98.0%; Pred. No. 2,6e-153;  
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Qy 1 MNRGVPFRHLLLVLTQALLPAATQGNKVVLGKGDVTELCTASQKKSIOFHMKNNOIK 60  
Db 1 MNRGVPFRHLLLVLTQALLPAATQGNKVVLGKGDVTELCTASQKKSIOFHMKNNOIK 60  
Qy 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
Qy 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVOCRSRPGKNIOGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVOCRSRPGKNIOGKTLVSQLELDQSG 180  
Qy 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEQVEFSPFLATFVEKLTGSGELMW 240  
Db 181 TWCTVLONQKQKVEFKIDIVLAFQKASSIYKKKEQVEFSPFLATFVEKLTGSGELMW 240  
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPOALPOYAGSGNLTLA 300  
Db 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPOALPOYAGSGNLTLA 300  
Qy 301 LEATGKLHQBVLVVMRATOLQKNTCEVWGPTSPKLMSTKLENKAKVSKREKPYWV 360  
Db 301 LEATGKLHQBVLVVMRATOLQKNTCEVWGPTSPKLMSTKLENKAKVSKREKPYWV 360  
Qy 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVAH--DPQLC 401  
Db 361 LNPAAGMWQCLLSDSGVLLSNIKVLPTWSTPVAH--DPQLC 401

RESULT 30

PCT-US95-00454-5  
Sequence 5, Application PC/TUS9500454

GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected  
TITLE OF INVENTION: Cells by chimeric CD4 Receptor-  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/00454  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/847,566  
FILING DATE: March 6, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:



LENGTH: 462 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-00454-5

Query Match 59.5%; Score 2032.5; DB 5; Length 462;  
Best Local Similarity 98.0%; Pred. No. 256-153;  
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVFRRLLVLTQALPAATQGNKVVGGKKDVTVELTCTASQKKSIOFHMKNSNQIK 60  
DB 1 MNRGVFRRLLVLTQALPAATQGNKVVGGKKDVTVELTCTASQKKSIOFHMKNSNQIK 60  
QY 61 ILGNQSFITKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQSFITKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240  
DB 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240  
QY 241 QABRASSKSWITFDLKNKEVSVKRVTQDPKIQMGKKLPLHLTLPALPOYAGSGNLTIA 300  
DB 241 QABRASSKSWITFDLKNKEVSVKRVTQDPKIQMGKKLPLHLTLPALPOYAGSGNLTIA 300  
QY 301 LEAKTGKHOEVNLYVMRATOLQKNLTCEWGPSTPKMLSLKLENKAIVSKREKPVWV 360  
DB 301 LEAKTGKHOEVNLYVMRATOLQKNLTCEWGPSTPKMLSLKLENKAIVSKREKPVWV 360  
QY 361 LNPEAGMOCCLSDSGQVLLSNIKVLPWTWSTPVPCAPPEPKSC 404  
DB 361 LNPEAGMOCCLSDSGQVLLSNIKVLPWTWSTPVVHA---DPOIC 401

## RESULT 31

US-08-284-391B-29  
Sequence 29, Application US/08284391B  
Patent No. 5851828  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-29

Query Match 59.4%; Score 2029; DB 2; Length 398;  
Best Local Similarity 100.0%; Pred. No. 4e-153;  
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVFRRLLVLTQALPAATQGNKVVGGKKDVTVELTCTASQKKSIOFHMKNSNQIK 60  
DB 1 MNRGVFRRLLVLTQALPAATQGNKVVGGKKDVTVELTCTASQKKSIOFHMKNSNQIK 60  
QY 61 ILGNQSFITKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQSFITKGPSKLNDRADSRSLMDQGNPFLIIKNKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240  
DB 181 TWCTVLQNKQKVEFKIDIVLAFOKASSIVYKKEGQVEFSFPLAFTVEKLTGSGELMW 240  
QY 241 QABRASSKSWITFDLKNKEVSVKRVTQDPKIQMGKKLPLHLTLPALPOYAGSGNLTIA 300  
DB 241 QABRASSKSWITFDLKNKEVSVKRVTQDPKIQMGKKLPLHLTLPALPOYAGSGNLTIA 300  
QY 301 LEAKTGKHOEVNLYVMRATOLQKNLTCEWGPSTPKMLSLKLENKAIVSKREKPVWV 360  
DB 301 LEAKTGKHOEVNLYVMRATOLQKNLTCEWGPSTPKMLSLKLENKAIVSKREKPVWV 360  
QY 361 LNPEAGMOCCLSDSGQVLLSNIKVLPWTWSTPV 394  
DB 361 LNPEAGMOCCLSDSGQVLLSNIKVLPWTWSTPV 394

## RESULT 32

US-09-218-950-29  
Sequence 29, Application US/09218950  
Patent No. 6284240  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-218-950-29

```

```

Query Match      59.4%; Score 2029; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 4e-153;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPRHLLVQLALPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
DB 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
QY 181 TWICTVLQONOKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWICTVLQONOKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKKLPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKKLPLHLTLPOALPOYAGSGNLTLA 300
QY 301 LEAKTGLHOEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSREKRVVW 360
DB 301 LEAKTGLHOEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSREKRVVW 360
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394

```

```

RESULT 33
US-08-466-368-4
; Sequence 4, Application US/08466368
; Patent No. 6035339
; GENERAL INFORMATION:
; APPLICANT: Maddon, Paul J.
; APPLICANT: Littman, Dan R.
; APPLICANT: Chess, Leonard

```

```

; APPLICANT: Axel, Richard
; APPLICANT: Weiss, Robin
; APPLICANT: McDougal, J. S.
; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN
; TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,368
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 24577-EI-B/JPM/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-368-4

```

```

Query Match      59.4%; Score 2029; DB 3; Length 458;
Best Local Similarity 100.0%; Pred. No. 4.9e-153;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPRHLLVQLALPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
DB 121 LVFGLTANSDTHLLOQGSLLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDG 180
QY 181 TWICTVLQONOKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWICTVLQONOKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKKLPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKKLPLHLTLPOALPOYAGSGNLTLA 300
QY 301 LEAKTGLHOEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSREKRVVW 360
DB 301 LEAKTGLHOEVNLVVMRATQLOKNTLCEVWGPTSPKLMSTLKENKEAVSREKRVVW 360
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPV 394

```

```

RESULT 34
US-08-328-500-9
; Sequence 9, Application US/08328500

```

Patent No. 6673896  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
APPLICANT: Axel, Richard  
APPLICANT: Sweet, Richard W.  
APPLICANT: Archos, James  
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/328,500  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/24577-CY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ. ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 457 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-328-500-9

Query Match 59.2%; Score 2021; DB 4; Length 457;  
Best Local Similarity 99.7%; Pred. No. 2.1e-152;

Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTALLPATQGNKVVGKKGPTVELTCTASOKKSIOFHMKSNQK 60  
DB 1 MNRGVPRHLLVLTALLPATQGNKVVGKKGPTVELTCTASOKKSIOFHMKSNQK 60  
QY 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNPFLIIKLIKIDSDTYICEVEDEKEEVOL 120  
DB 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNPFLIIKLIKIDSDTYICEVEDEKEEVOL 120  
QY 121 LVFGLTANSPTLHLLQGSILTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPTLHLLQGSILTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTVLTQNKQKVEFKIDIVVLAFOKASSIVYKKGSEQVEFPLAFYVEKLTGSGELMW 240  
DB 181 TWCTVLTQNKQKVEFKIDIVVLAFOKASSIVYKKGSEQVEFPLAFYVEKLTGSGELMW 240  
QY 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOALPOYAGSGMLTTLA 300  
DB 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOALPOYAGSGMLTTLA 300  
QY 301 LEAKTGKLGQEVNVLVVRATQLOKNLTCEVWGSPSPKMLSLKENKAKVSRREKPVVW 360  
DB 301 LEAKTGKLGQEVNVLVVRATQLOKNLTCEVWGSPSPKMLSLKENKAKVSRREKPVVW 360  
QY 361 LNPBAGMOCCLSDSGVLLSNIKULPTWSTPV 394  
DB 361 LNPBAGMOCCLSDSGVLLSNIKULPTWSTPV 394

RESULT 35

US-08-236-311-1  
Sequence 1, Application US/08236311

Patent No. 5565335  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 444P1C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ. ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-236-311-1

Query Match 59.1%; Score 2017; DB 1; Length 402;  
Best Local Similarity 99.7%; Pred. No. 3.6e-152;

Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTALLPATQGNKVVGKKGPTVELTCTASOKKSIOFHMKSNQK 60  
DB 1 MNRGVPRHLLVLTALLPATQGNKVVGKKGPTVELTCTASOKKSIOFHMKSNQK 60  
QY 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNPFLIIKLIKIDSDTYICEVEDEKEEVOL 120  
DB 61 ILGNQSFLLTKGSKLNDRADSRSLMDQGNPFLIIKLIKIDSDTYICEVEDEKEEVOL 120  
QY 121 LVFGLTANSPTLHLLQGSILTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSPTLHLLQGSILTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTVLTQNKQKVEFKIDIVVLAFOKASSIVYKKGSEQVEFPLAFYVEKLTGSGELMW 240  
DB 181 TWCTVLTQNKQKVEFKIDIVVLAFOKASSIVYKKGSEQVEFPLAFYVEKLTGSGELMW 240  
QY 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOALPOYAGSGMLTTLA 300  
DB 241 QARRASSSKSWITFDLKNKEVSVKRVYQDPKLGKGLPLHLTLPOALPOYAGSGMLTTLA 300

Db 241 QABASSSKSMITFDLKNKEVSVRKVTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLA 300  
Qy 301 LEATGKLGHOEVNLVWRATQLOKNTCEVWGPTSPKLMSTLKENKEAKVSKREKPYWV 360  
Db 301 LEATGKLGHOEVNLVWRATQLOKNTCEVWGPTSPKLMSTLKENKEAKVSKREKAVWV 360  
Qy 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393  
Db 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393

## RESULT 36

US-08-457-918-1  
Sequence 1, Application US/08457918  
Patent No. 6117655  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variance  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 Inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-457-918-1

Query March 59.1%; Score 2017; DB 3; Length 402;  
Best Local Similarity 99.7%; Pred. No. 3.6e-152;

Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MNRGVPRHLLVLQALLPATQGNKVVLRKSGDTEVELTCTASQKSIOPHMKNSQIK 60  
Db 1 MNRGVPRHLLVLQALLPATQGNKVVLRKSGDTEVELTCTASQKSIOPHMKNSQIK 60

Qy 61 ILGNQSGFLTKGSPKLNDRADSRRLMDQGNFPLIINKLTIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILGNQSGFLTKGSPKLNDRADSRRLMDQGNFPLIINKLTIEDSDTYICEVEDQKEEVOL 120  
Qy 121 LVFGLTANSDTHLLQGGSLTLLTSPGSSPSVQCRSPRGKNIQGGKTLISVSOLELQDSG 180  
Db 121 LVFGLTANSDTHLLQGGSLTLLTSPGSSPSVQCRSPRGKNIQGGKTLISVSOLELQDSG 180  
Qy 181 TWCTTVLQNKQKVEFKDIVLAFQKASSIVYKKEGQVEFSPPLATVEKLTGSGGLMW 240  
Db 181 TWCTTVLQNKQKVEFKDIVLAFQKASSIVYKKEGQVEFSPPLATVEKLTGSGGLMW 240  
Qy 241 QABASSSKSMITFDLKNKEVSVRKVTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLA 300  
Db 241 QABASSSKSMITFDLKNKEVSVRKVTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLA 300  
Qy 301 LEATGKLGHOEVNLVWRATQLOKNTCEVWGPTSPKLMSTLKENKEAKVSKREKPYWV 360  
Db 301 LEATGKLGHOEVNLVWRATQLOKNTCEVWGPTSPKLMSTLKENKEAKVSKREKAVWV 360  
Qy 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393  
Db 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTP 393

## RESULT 37

US-09-039-555B-15  
Sequence 15, Application US/09039555B  
Patent No. 6033856  
GENERAL INFORMATION:  
APPLICANT: Koerner, Kathrin  
APPLICANT: Mueller, Rolf  
TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,555B  
FILING DATE: 16-MAR-1998  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 19710643.9  
FILING DATE: 14-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 016779/0131  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 458 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-039-555B-15

Query March 59.1%; Score 2017; DB 3; Length 402;  
Best Local Similarity 99.7%; Pred. No. 3.6e-152;

Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MNRGVPRHLLVLQALLPATQGNKVVLRKSGDTEVELTCTASQKSIOPHMKNSQIK 60  
Db 1 MNRGVPRHLLVLQALLPATQGNKVVLRKSGDTEVELTCTASQKSIOPHMKNSQIK 60

Query Match 59.0%; Score 2015; DB 3; Length 458;  
Best Local Similarity 99.5%; Pred. No. 6.3e-152;  
Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKKGDVETLCTASOKKSIQFHMKNNOIK 60  
DB 1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKKGDVETLCTASOKKSIQFHMKNNOIK 60  
QY 61 IIGNGSSFLTKGSPSKNDRAADRSRSLMDQGNPFLIIRNKLIKEDSDTYICEVEDQKEVOL 120  
DB 61 IIGNGSSFLTKGSPSKNDRAADRSRSLMDQGNPFLIIRNKLIKEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDBTHLLOQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDBTHLLOQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTVLOQKKEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
DB 181 TWCTVLOQKKEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
QY 241 QAEARSSSKSWITFDLKNKEVSVKRYTODPKLQMGKLLPLHLTLPQALPOYAGSGNLTLA 300  
DB 241 QAEARSSSKSWITFDLKNKEVSVKRYTODPKLQMGKLLPLHLTLPQALPOYAGSGNLTLA 300  
QY 301 LEAKTKLHOEVNLVWBAATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360  
DB 301 LEAKTKLHOEVNLVWBAATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360  
QY 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTPV 394  
DB 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTPV 394

RESULT 38  
US-09-517-605-3  
; Sequence 3, Application US/09517605  
; Patent No. 6391567  
; GENERAL INFORMATION:  
; APPLICANT: Littman, Dan R.  
; APPLICANT: Kwon, Douglas S.  
; APPLICANT: van Kooyk, Yvette  
; APPLICANT: Geijtenbeek, Theo  
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO  
; FILE REFERENCE: 1049-1-017  
; CURRENT APPLICATION NUMBER: US/09/517,605  
; CURRENT FILING DATE: 2000-03-02  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-517-605-3

Query Match 58.8%; Score 2007; DB 4; Length 458;  
Best Local Similarity 99.2%; Pred. No. 2.7e-151;  
Matches 391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKKGDVETLCTASOKKSIQFHMKNNOIK 60  
DB 1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKKGDVETLCTASOKKSIQFHMKNNOIK 60  
QY 61 IIGNGSSFLTKGSPSKNDRAADRSRSLMDQGNPFLIIRNKLIKEDSDTYICEVEDQKEVOL 120  
DB 61 IIGNGSSFLTKGSPSKNDRAADRSRSLMDQGNPFLIIRNKLIKEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDBTHLLOQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDBTHLLOQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWCTVLOQKKEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
DB 181 TWCTVLOQKKEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240

DB 181 TWCTVLOQKKEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
QY 241 QAEARSSSKSWITFDLKNKEVSVKRYTODPKLQMGKLLPLHLTLPQALPOYAGSGNLTLA 300  
DB 241 QAEARSSSKSWITFDLKNKEVSVKRYTODPKLQMGKLLPLHLTLPQALPOYAGSGNLTLA 300  
QY 301 LEAKTKLHOEVNLVWBAATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360  
DB 301 LEAKTKLHOEVNLVWBAATOLQKNLTCEVWGPTSPKMLSLKLNKEAKVSKREKPYWV 360  
QY 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTPV 394  
DB 361 LNPEAGMWOCCLSDSGVLLSNIKVLPTWSTPV 394

RESULT 39  
US-08-466-368-2  
; Sequence 2, Application US/08466368  
; Patent No. 6093539  
; GENERAL INFORMATION:  
; APPLICANT: Madden, Paul J.  
; APPLICANT: Littman, Dan R.  
; APPLICANT: Chess, Leonard  
; APPLICANT: Axel, Richard  
; APPLICANT: Weiss, Robin  
; APPLICANT: McDougall, J. S.  
; TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,368  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 24577-EI-B/JPW/AKC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 394 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-466-368-2

Query Match 58.6%; Score 2001; DB 3; Length 394;  
Best Local Similarity 98.7%; Pred. No. 6.6e-151;  
Matches 389; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKKGDVETLCTASOKKSIQFHMKNNOIK 60  
DB 1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKKGDVETLCTASOKKSIQFHMKNNOIK 60  
QY 61 IIGNGSSFLTKGSPSKNDRAADRSRSLMDQGNPFLIIRNKLIKEDSDTYICEVEDQKEVOL 120  
DB 61 IIGNGSSFLTKGSPSKNDRAADRSRSLMDQGNPFLIIRNKLIKEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDBTHLLOQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDBTHLLOQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180

DB 121 LVFGLTANSDTHLQGGSLTTLTLESPGSSSPVOCRRPRGNIOGKTLVSQLELDPSG 180  
OY 181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGEVDFSPFLAFTVEKLTSSGELMW 240  
DB 181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGEVDFSPFLAFTVEKLTSSGELMW 240  
OY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300  
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300  
OY 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMSTLKLENKAKVSKREKPYV 360  
DB 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMSTLKLENKAKVSKREKPYV 360  
OY 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394  
DB 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394

## RESULT 40

US-08-328-500-2  
Sequence 2, Application US/08328500

Patent No. 6673896  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
APPLICANT: Axel, Richard  
APPLICANT: Sweet, Richard W.  
APPLICANT: Athos, James  
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/328,500  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/24577-CY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 394 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-328-500-2

Query Match 58.6%; Score 2001; DB 4; Length 394;  
Best Local Similarity 98.7%; Pred. No. 6.6e-151;  
Matches 389; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 MNRGVPRHLLVQLALLPATOGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPATOGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
OY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120  
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120

OY 121 LVFGLTANSDTHLQGGSLTTLTLESPGSSSPVOCRRPRGNIOGKTLVSQLELDPSG 180  
DB 121 LVFGLTANSDTHLQGGSLTTLTLESPGSSSPVOCRRPRGNIOGKTLVSQLELDPSG 180  
OY 181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGEVDFSPFLAFTVEKLTSSGELMW 240  
DB 181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGEVDFSPFLAFTVEKLTSSGELMW 240  
OY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300  
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300  
OY 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMSTLKLENKAKVSKREKPYV 360  
DB 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMSTLKLENKAKVSKREKPYV 360  
OY 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394  
DB 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394

## RESULT 41

5223394-7

Patent No. 5223394  
APPLICANT: WALLNER, BARBARA  
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
LINKAGE SIGNAL SEQUENCE  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/335,688  
FILING DATE: 10-APR-1989  
SEQ ID NO: 7  
LENGTH: 458

5223394-7

Query Match 58.3%; Score 1989; DB 6; Length 458;  
Best Local Similarity 98.5%; Pred. No. 7.4e-150;  
Matches 388; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 MNRGVPRHLLVQLALLPATOGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPATOGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
OY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120  
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120  
OY 121 LVFGLTANSDTHLQGGSLTTLTLESPGSSSPVOCRRPRGNIOGKTLVSQLELDPSG 180  
DB 121 LVFGLTANSDTHLQGGSLTTLTLESPGSSSPVOCRRPRGNIOGKTLVSQLELDPSG 180  
OY 181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGEVDFSPFLAFTVEKLTSSGELMW 240  
DB 181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGEVDFSPFLAFTVEKLTSSGELMW 240  
OY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300  
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPOALPOYAGSGNLTIA 300  
OY 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMSTLKLENKAKVSKREKPYV 360  
DB 301 LEATGKLGHOEVNLVVMRATOLQKULTCCEVWGPTSPKLMSTLKLENKAKVSKREKPYV 360  
OY 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394  
DB 361 LNPEAGMWOCCLSDSGOVLLESNIKVLPTWSTPV 394

RESULT 42  
5223418-2  
Patent No. 5223418

APPLICANT: ARCURI, EDWARD J.;BRAUNER, MARY E.;DONOVAN, MARY  
J.;GERBER, ROBERT G.;KELLER, JOHN A.  
TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF  
HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS  
NUMBER OF SEQUENCES: 2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/589,979  
FILING DATE: 28-SEP-1990  
SEQ ID NO:2  
LENGTH: 394  
5223418-2

Query Match 57.1%; Score 1951; DB 6; Length 394;  
Best Local Similarity 97.5%; Pred. No. 6,2e-147;  
Matches 384; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAAQGNKVKVGGKDTVELCTASQKKSIOFHKNSNOIK 60  
DB 1 MNRGVPFRHLVLVQLALPAAQGNKVKVGGKDTVELCTASQKKSIOFHKNSNOIK 60  
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIITKNLIEDSDTYICEVEDQKEVOL 120  
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIITKNLIEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSPDTHLLOGSLTLTLSPSPSSPVQCSPPKKNIOGGKTLISVQLELDQSG 180  
DB 121 LVFGLTANSPDTHLLOGSLTLTLSPSPSSPVQCSPPKKNIOGGKTLISVQLELDQSG 180  
QY 181 TWTCTVLQNGKVEFKIDIVLAFQKASIVYKKEGEQVEFSFPLAFTVEKLTGSGELMW 240  
DB 181 TWTCTVLQNGKVEFKIDIVLAFQKASIVYKKEGEQVEFSFPLAFTVEKLTGSGELMW 240  
QY 241 QABRASSKSWITFDLKNKSVKRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTLA 300  
DB 241 QABRASSKSWITFDLKNKSVKRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTLA 300  
QY 301 LKAKTGKHOEVNLVVRATQLOKNTLCEVWGPTSPKMLSLKENKAKVSKREKAVNY 360  
DB 301 LKAKTGKHOEVNLVVRATQLOKNTLCEVWGPTSPKMLSLKENKAKVSKREKAVNY 360  
QY 361 LNPBAGMOCCLSDSGQVLLSNIKVLPTWSTPV 394  
DB 361 LNPBAGMOCCLSDSGQVLLSNIKVLPTWSTPV 394

RESULT 43  
US-08-236-311-4  
Sequence 4, Application US/08236311  
Patent No. 5565335  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adheson Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 444P1C2  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 434 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-236-311-4

Query Match 55.8%; Score 1904; DB 1; Length 434;  
Best Local Similarity 99.7%; Pred. No. 3.9e-143;  
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 QGNKVVLAGKKGDTVELCTASQKKSIOFHKNSNOIKILGNQGSFLTKGPSKLNDRADR 83  
DB 56 QGNKVVLAGKKGDTVELCTASQKKSIOFHKNSNOIKILGNQGSFLTKGPSKLNDRADR 115  
QY 84 RSLMDQGNFPLIITKNLIEDSDTYICEVEDQKEVOLLVFGLTANSPDTHLLOGSLTLTL 143  
DB 116 RSLMDQGNFPLIITKNLIEDSDTYICEVEDQKEVOLLVFGLTANSPDTHLLOGSLTLTL 175  
QY 144 ESPSPSSPVQCSPPKKNIOGGKTLISVQLELDQSGTCTVLQNGKVEFKIDIVLVA 203  
DB 176 ESPSPSSPVQCSPPKKNIOGGKTLISVQLELDQSGTCTVLQNGKVEFKIDIVLVA 235  
QY 204 FQKASIVYKKEGEQVEFSFPLAFTVEKLTGSGELMWAQBRASSKSWITFDLKNKESV 263  
DB 236 FQKASIVYKKEGEQVEFSFPLAFTVEKLTGSGELMWAQBRASSKSWITFDLKNKESV 295  
QY 264 KRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTALAKTGKHOEVNLVVRATQLO 323  
DB 296 KRVTDQPKLQMGKPLHLTLPOALPOYAGSGNLTALAKTGKHOEVNLVVRATQLO 355  
QY 324 KNLTCVWGPTSPKMLSLKENKAKVSKREKAVVNLNPBAGMOCCLSDSGQVLLSN 383  
DB 356 KNLTCVWGPTSPKMLSLKENKAKVSKREKAVVNLNPBAGMOCCLSDSGQVLLSN 415  
QY 384 IKVLPTWSTP 393  
DB 416 IKVLPTWSTP 425

RESULT 44  
US-08-457-918-4  
Sequence 4, Application US/08457918  
Patent No. 6117655  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adheson Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080



```

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genetech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,918
FILING DATE: 1-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/236311
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Kubinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0444P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 434 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-457-918-4

Query Match: 55.8%; Score 1904; DB 3; Length 434;
Best Local Similarity 99.7%; Pred. No. 3.9e-143;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

24 QGNVVLGKKGDVLTCTASQKKSIOFHWKNSNOIKILGNQGSFLTKGPSKLNDRADSR 83
DB 56 QGNVVLGKKGDVLTCTASQKKSIOFHWKNSNOIKILGNQGSFLTKGPSKLNDRADSR 115
QY RSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSFLTTL 143
DB 116 RSLMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSFLTTL 175
QY 144 ESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLIQNKQVEFKIDIVLA 203
DB 176 ESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLIQNKQVEFKIDIVLA 235
QY 204 FQKASSIVYKKEGGEVERSPFLATVEKLTGSGGELMWAERASSSKSWITFDLKNKEVS 263
DB 236 FQKASSIVYKKEGGEVERSPFLATVEKLTGSGGELMWAERASSSKSWITFDLKNKEVS 295
QY 264 KRVTPDLQMGKLPPLHTLTPQALPOYAGSGNLTALAEATGKLHGEVNLVWMRATOLQ 323
DB 296 KRVTPDLQMGKLPPLHTLTPQALPOYAGSGNLTALAEATGKLHGEVNLVWMRATOLQ 355
QY 324 KNLTCGEVWGPSPKMLSLKLENKEAKVSKREKPVWVNLPEAGMOCILSDSGVLLBSN 383
DB 356 KNLTCGEVWGPSPKMLSLKLENKEAKVSKREKPVWVNLPEAGMOCILSDSGVLLBSN 415
QY 384 IKVLPWTSTP 393
DB 416 IKVLPWTSTP 425

```

RESULT 45

```

US-08-867-149-1
Sequence 1, Application US/08867149
Patent No. 5912176
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Antibody Against a Host Cell
TITLE OF INVENTION: Antigen Complex for Pre- and Post-Exposure
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maria C.H. Lin
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,149
FILING DATE: 28-Feb-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/657,149
FILING DATE: 03-June-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4145
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849
TELEFAX: (212)415-8745
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 433 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-149-1

Query Match: 55.4%; Score 1892; DB 2; Length 433;
Best Local Similarity 99.5%; Pred. No. 3.5e-142;
Matches 367; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

26 NKVVLGKKGDVLTCTASQKKSIOFHWKNSNOIKILGNQGSFLTKGPSKLNDRADSR 85
DB 1 NKVVLGKKGDVLTCTASQKKSIOFHWKNSNOIKILGNQGSFLTKGPSKLNDRADSR 60
QY 86 LMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSFLTTL 145
DB 61 LMDQGNFPLIIKNIKIEDSDTYICEVEDQKEEVQLVFGLTANSPTHLQGSFLTTL 120
QY 146 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLIQNKQVEFKIDIVLAQ 205
DB 121 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLIQNKQVEFKIDIVLAQ 180
QY 206 KASSIVYKKEGGEVERSPFLATVEKLTGSGGELMWAERASSSKSWITFDLKNKEVS 265
DB 181 KASSIVYKKEGGEVERSPFLATVEKLTGSGGELMWAERASSSKSWITFDLKNKEVS 240
QY 266 VTQDPKLGKGLPPLHTLTPQALPOYAGSGNLTALAEATGKLHGEVNLVWMRATOLQ 325
DB 241 VTQDPKLGKGLPPLHTLTPQALPOYAGSGNLTALAEATGKLHGEVNLVWMRATOLQ 300
QY 326 LTCGEVWGPSPKMLSLKLENKEAKVSKREKPVWVNLPEAGMOCILSDSGVLLBSN 385
DB 301 LTCGEVWGPSPKMLSLKLENKEAKVSKREKPVWVNLPEAGMOCILSDSGVLLBSN 360

```

Db 182 ASSIVYKKEGEQVERSFPLAFTVEKLTGGSELMWQARASSKSWIIFDLKKNKEVSVRK 241

QY 267 TQDPEKQWKKLPLHLTLPOALPOYAGSGNLTLLAEKTKLHQBVLVWMRATQLOKNL 326  
 DB 242 TQDPEKQWKKLPLHLTLPOALPOYAGSGNLTLLAEKTKLHQBVLVWMRATQLOKNL 301  
 QY 327 TCEVWGPTSPKMLSLKLENKAVSKREKPVWVLPBAGMOCCLSDSGVLLIESNIKV 386  
 DB 302 TCEVWGPTSPKMLSLKLENKAVSKREKPVWVLPBAGMOCCLSDSGVLLIESNIKV 361  
 QY 387 LPTWSTPV 394  
 DB 362 LPTWSTPV 369

RESULT 48  
 5171838-13  
 Patent No. 5171838  
 APPLICANT: CHIBA, YUKINOBU  
 TITLE OF INVENTION: LEU3A BINDING PEPTIDES  
 NUMBER OF SEQUENCES: 24  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/526,921  
 FILING DATE: 22-MAY-1990  
 SEQ ID NO:13:  
 LENGTH: 433  
 5171838-13

Query Match 49.6%; Score 1695; DB 6; Length 433;  
 Best Local Similarity 91.2%; Pred. No. 1,66-126;  
 Matches 342; Conservative 6; Mismatches 13; Indels 14; Gaps 4;

QY 27 KVVLGKGGDTVELCTASOKKSIQFHWKNSNOIKILGQSGFLTKGPKLNDRAISR 85  
 DB 2 KVVLGKGGDTVELCTASOKKSIQFHWKNSNOIKILGQSGFLTKGPKLNDRAISR 61  
 QY 86 ----LWDQGNPLIIKNLKIED-SDTYICEVEDQKEVQLVFGLTANSDTHLLOQSL 139  
 DB 62 NQIKLGKXGSF-LTKGPKLNDRAISR-----RRSEVQLVFGLTANSDTHLLOQSL 114  
 QY 140 TLTIESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDSGTWCTVYLQNKVEFKIDI 199  
 DB 115 TLTIESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDSGTWCTVYLQNKVEFKIDI 174  
 QY 200 VLAFOKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMWQABRASSSKSWITFDLKNK 259  
 DB 175 VLAFOKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMWQABRASSSKSWITFDLKNK 234  
 QY 260 EVSVKRVYQDPEKQWKKLPLHLTLPOALPOYAGSGNLTLLAEKTKLHQBVLVWMRA 319  
 DB 235 EVSVKRVYQDPEKQWKKLPLHLTLPOALPOYAGSGNLTLLAEKTKLHQBVLVWMRA 294  
 QY 320 TQLOKNLTCEVWGPTSPKMLSLKLENKAVSKREKPVWVLPBAGMOCCLSDSGVYL 379  
 DB 295 TQLOKNLTCEVWGPTSPKMLSLKLENKAVSKREKPVWVLPBAGMOCCLSDSGVYL 354  
 QY 380 LESNIKVLPTWSTPV 394  
 DB 355 LESNIKVLPTWSTPV 369

RESULT 49  
 US-08-630-172-17  
 Sequence 17, Application US/08630172  
 Patent No. 6060054  
 GENERAL INFORMATION:  
 APPLICANT: Staerz, Uwe  
 TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheridan Rose & McIntosh  
 STREET: 1700 Lincoln Street, 35th Floor  
 CITY: Denver  
 STATE: Colorado

COUNTRY: U.S.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/630,172  
 FILING DATE:  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Connell, Gary J.  
 REGISTRATION NUMBER: 32,020  
 REFERENCE/DOCKET NUMBER: 2879-36  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 410 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-630-172-17

Query Match 46.4%; Score 1585.5; DB 3; Length 410;  
 Best Local Similarity 54.2%; Pred. No. 7,36-118;  
 Matches 325; Conservative 33; Mismatches 51; Indels 191; Gaps 6;

QY 26 NKVVLGKGGDTVELCTASOKKSIQFHWKNSNOIKILGQSGFLTKGPKLNDRAISR 85  
 DB 1 NKVVLGKGGDTVELCTASOKKSIQFHWKNSNOIKILGQSGFLTKGPKLNDRAISR 60  
 QY 86 LWDQGNPLIIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDTHLLOQSLTLTLES 145  
 DB 61 LWDQGNPLIIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDTHLLOQSLTLTLES 120  
 QY 146 PPGSSPVQCRSPRGKNIQGGKTLVSQLELDSGTWCTVYLQNKVEFKIDI VLAFO 205  
 DB 121 PPGSSPVQCRSPRGKNIQGGKTLVSQLELDSGTWCTVYLQNKVEFKIDI VLAFO 178  
 QY 206 KASSIVYKKEGEVFEFPLAFTVEKLTGSGELMWQABRASSSKSWITFDLKNKVEYKR 265  
 DB 179 ----- 178  
 QY 266 VTQDPEKQWKKLPLHLTLPOALPOYAGSGNLTLLAEKTKLHQBVLVWMRATQLOKN 325  
 DB 179 ---EPR----- 181  
 QY 326 LTCVWGPTSPKMLSLKLENKAVSKREKPVWVLPBAGMOCCLSDSGVLLIESNIK 385  
 DB 182 -----GPT----- 186  
 QY 386 VLPWSTPVPCAPAPPEKSCDKHTCPQLDGPSTVLEPPKPKDTLMTSRTPEVTCVVVDV 445  
 DB 187 PCP-----PCPCPA-----PVLGSPSTVLEPPKPKDTLMTSRTPEVTCVVVDV 230  
 QY 446 SHEDPEVFMVYDGVENHAKTPREDOYNSYRVVSVTLVHODVLNGEKYCKSNK 505  
 DB 231 SEDDPDVQISMFVNNAVHTAQOTHTREDNSRKRVSALPIQHDQMSGKPEFKYNNK 290  
 QY 506 ALPAPIKTSKAKQPREPOVYTLTPSRDELTKQVSLTCLVKGYPYSIAVWESNQ 565  
 DB 291 DLPAPIRITSKPKGSVRAPQVYVLPPEP-DEMTKQVTLTQMTDFPEDIYVEMTNNGK 349  
 QY 566 PENNYKTPPVLDSDGSFPLYSKLTVDKSNQGNVSCSVMBEALHNHYTQSLSPG 625  
 DB 350 TELNYKTEPVLDSDGSYFMYSKLRVEKKNWVERNSYCSVHGLHNHHTTYSFSHTPG 409

RESULT 50  
 US-09-375-419-17

```

; Sequence 17, Application US/09375419
; Patent No. 6264950
; GENERAL INFORMATION:
; APPLICANT: Staerz, Uwe
; TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
; TITLE OF INVENTION: LYMPHOCYTE VERO
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, 35th Floor
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/375,419
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,172
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-375-419-17

Query Match      46.4%; Score 1585.5; DB 3; Length 410;
Best Local Similarity 54.2%; Pred. No. 7,3e-118;
Matches 32; Conservative 33; Mismatches 51; Indels 191; Gaps 6;

QY      26 NKVVLGGKGDVTELTCTASQKKSIOFHWNKSNQIKITGNQGSFLTKGPSKLNDRADSRRS 85
DB      1 NKVVLGGKGDVTELTCTASQKKSIOFHWNKSNQIKITGNQGSFLTKGPSKLNDRADSRRS 60

QY      86 LMOGNEPLTIKULKIEDSDTYICEVEDQKEVQLVFGLTANSPTHTLLOGSLTLTLES 145
DB      61 LMOGNEPLTIKULKIEDSDTYICEVEDQKEVQLVFGLTANSPTHTLLOGSLTLTLES 120

QY      146 PGGSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLONQKVEFKIDIVLAFQ 205
DB      121 PGGSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTCTVLONQKVEFKIDIVLAFQ 178

QY      206 KASSIVYKKEGEVFSFPLAFVTEKLTGSGELMWAQERASSKSWITFDLKNKEVSVKR 265
DB      179 ----- 178

QY      266 VTQDPKIQMGKULPLHLTLPOALPOVAGSGNLTALAEAKTGKLAHQEVNLVVMARATQLOKN 325
DB      179 ----- 181

QY      326 LTCGVMPGTPSKLMLSLKENKEAKVSRKRPVWVLNPEAGMOCCLSDSGQVLTLESNIK 385
DB      182 -----IK 186

QY      386 VLPTWSTPVPCEPAEPKSCDKTHTCCPELLGSPSYFLFPKPKDTLMISRTPEVATCVVVDV 445
DB      187 PCP-----PCKCPA-----PMLGGSPSVFIRPPPKIKDVLMLISLSPITVTCVVVDV 230

```

```

QY      446 SHEDEPKENMYVDGVEYENAKTKPREBOYNSTYRVSVLTVLHODMLNGKEYCKCVSNK 505
DB      231 SEDDEPDQISWVFVNVNEVHTAQTQREDYNSRLKAVSLPQIHDQMGSKCKKCVNNK 290

QY      506 ALPAPIEKTIKAKQOPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNQ 565
DB      291 DLPAPIERTISKPKSVRAFYVYVLPPE-EWTKKQVTLTCVTDPMFEDIYVEMTNNGK 349

QY      566 PENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGVNFSGVMEALHNYTQKSLSLSPG 625
DB      350 TELNYKNTPEVLDSDGSFYFWSKLVKKNWERNRNSYCSVVEGLHNNHTTKSFSTRG 409

RESULT 51
5223394-11
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO:11:
; LENGTH: 318
5223394-11

Query Match      40.1%; Score 1368; DB 6; Length 318;
Best Local Similarity 93.4%; Pred. No. 9,7e-101;
Matches 268; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

QY      1 NMRGVFRRLLLVLTQALLPAATQGNKVVLTGKKGDTVELTCTASQKKSIOFHWNKSNQIK 60
DB      1 NMRGVFRRLLLVLTQALLPAATQGNKVVLTGKKGDTVELTCTASQKKSIOFHWNKSNQIK 60

QY      61 ILGNQGSFLTKGPSKLNDRADSRRSIMDQGNFPLTIKULKIEDSDTYICEVEDQKEVQL 120
DB      61 ILGNQGSFLTKGPSKLNDRADSRRSIMDQGNFPLTIKULKIEDSDTYICEVEDQKEVQL 120

QY      121 LVFGLTANSPTHTLLOGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDG 180
DB      121 LVFGLTANSPTHTLLOGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDG 180

QY      181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVFSFPLAFVTEKLTGSGELMW 240
DB      181 TWTCTVLONQKVEFKIDIVLAFQKASSIVYKKEGEVFSFPLAFVTEKLTGSGELMW 240

QY      241 QABRASSKSWITFDLKNKEVSVKRVTDPKIQMGKULPLHLTLPOA 287
DB      241 QABRASSKSWITFDLKNKEVSVKRVISNPLNTTSSIIITTCIPSS 287

RESULT 52
5223394-9
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO:9:
; LENGTH: 295
5223394-9

Query Match      39.9%; Score 1363; DB 6; Length 295;
Best Local Similarity 99.3%; Pred. No. 2,2e-100;
Matches 266; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 NMRGVFRRLLLVLTQALLPAATQGNKVVLTGKKGDTVELTCTASQKKSIOFHWNKSNQIK 60

```

```

Db 1 MNRGPFPHLLLVQLALLPATQKQVVLGKGDVLELTCTASQKSIQFHMKNQIK 60
Qy 61 ILNGSGFLTKGPKLNDNRADSRSLMNGNFPILIKLKLEDDTYICEVEDKEEVOL 120
Db 61 ILNGSGFLTKGPKLNDNRADSRSLMNGNFPILIKLKLEDDTYICEVEDKEEVOL 120
Qy 121 LVFLGTLNSDTHLLOQSGSLTLTLESPGSSPVOCSPRGKNIQSGKTLVSQLELODSG 180
Db 121 LVFLGTLNSDTHLLOQSGSLTLTLESPGSSPVOCSPRGKNIQSGKTLVSQLELODSG 180
Qy 181 TWCTVLOQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
Db 181 TWCTVLOQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
Qy 241 QAERASSSKSWITFDLKNKEVSVKRVYQ 268
Db 241 QAERASSSKSWITFDLKNKEVSVKRVYR 268

```

## RESULT 53

US-08-284-391B-33

Sequence 33, Application US/08284391B

Patent No. 5851828

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Banapour, Babak

APPLICANT: Romeo, Charles

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark &amp; Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391B

FILING DATE: 02-AUG-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX:

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 254 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-284-391B-33

Query Match 39.2%; Score 1338.5; DB 2; Length 254;

Best Local Similarity 98.0%; Pred. No. 1.6e-98;  
Matches 249; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

```

Qy 400 EPKSCDTHHC-----PELLGSPVFLFPPKPKDTLMISRTPEYTCVVVDVSHEDPEVKF 454
Db 1 EPKSCDTHHCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEYTCVVVDVSHEDPEVKF 60
Qy 455 NMVYDGVVHNAKPKPEEOYNSTYRVVSVLTVLAHDMLNKEKCKVSNKALPAPIEKT 514
Db 61 NMVYDGVVHNAKPKPEEOYNSTYRVVSVLTVLAHDMLNKEKCKVSNKALPAPIEKT 120
Qy 515 ISKAKGPREPQVYTLPPSRDELTKQVSLTCLVKGFPSPDIAVEMESNQGPENNYKTP 574
Db 121 ISKAKGPREPQVYTLPPSRDELTKQVSLTCLVKGFPSPDIAVEMESNQGPENNYKTP 180
Qy 575 PVLDSGDSFPLYSKLTVDKSRWQGNVFGSGVHMEALHNYTKSLSLSPQLDFTCAE 634
Db 181 PVLDSGDSFPLYSKLTVDKSRWQGNVFGSGVHMEALHNYTKSLSLSPQLDFTCAE 240
Qy 635 AODGELDGLWTTDP 648
Db 241 AODGELDGLWTTDP 254

```

## RESULT 54

US-09-218-950-33

Sequence 33, Application US/09218950

Patent No. 6284240

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Banapour, Babak

APPLICANT: Romeo, Charles

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark &amp; Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/218,950

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391

FILING DATE: 02-AUG-1994

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX:

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 254 amino acids

TYPE: amino acid

Query Match 39.2%; Score 1338.5; DB 2; Length 254;

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-33

Query Match 39.2%; Score 1338.5; DB 3; Length 254;  
Best Local Similarity 98.0%; Pred. No. 1.6e-98;  
Matches 249; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 400 EPKSCDKHTC-----PELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 454  
DB 1 EPKSCDKHTCPCPCAPPELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
QY 455 NMVVDGVEVHNAKTKREEQDYNSTYYRVSVLTVLDHQMNGKEVKCKVKSKALPAPIEKT 514  
DB 61 NMVVDGVEVHNAKTKREEQDYNSTYYRVSVLTVLDHQMNGKEVKCKVKSKALPAPIEKT 120  
QY 515 ISKAGQPREPQVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 574  
DB 121 ISKAGQPREPQVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 180  
QY 575 PVLDSGSPFLYSKLTVDKSRMQGVNFCSCVMHEALHNHYTQKSLSLSPG 634  
DB 181 PVLDSGSPFLYSKLTVDKSRMQGVNFCSCVMHEALHNHYTQKSLSLSPGLDDETCAE 240  
QY 635 AADGELDGLMTTDP 648  
DB 241 AADGELDGLMTTDP 254

## RESULT 55

US-09-313-942-8  
Sequence 8, Application US/09313942  
Patent No. 6472179  
GENERAL INFORMATION:  
APPLICANT: REGENERON PHARMACEUTICALS, INC.  
TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING  
TITLE OF INVENTION: AND USING  
FILE REFERENCE: REG 203-A  
CURRENT APPLICATION NUMBER: US/09/313,942  
CURRENT FILING DATE: 1999-05-19  
PRIOR APPLICATION NUMBER: 09/313,942  
PRIOR FILING DATE: 1999-05-19  
PRIOR APPLICATION NUMBER: 60/101,858  
PRIOR FILING DATE: 1998-09-25  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 592  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-313-942-8

Query Match 37.3%; Score 1275; DB 4; Length 592;  
Best Local Similarity 47.3%; Pred. No. 5.9e-93;  
Matches 308; Conservative 53; Mismatches 160; Indels 130; Gaps 19;

QY 20 PAATQGNKVVYLGKKDTEVLTCTASQ-KKSIQPHMKNNOIKILGNQSFITKGPCKLND 78  
DB 26 PAQEVARGVLTSLPGDSVLTLCPEVEDNATVHM-----VLAKPA----- 66  
QY 79 RADSRRLMDQGNFPLIKULKIEDSDTYICE-----VEDQKEVQLVYFGLT 126  
DB 67 -ASHSPRRMAGMGRRLRLRSVOLHDSGYSCTYRAGPAGTVHLLVDPPEEPDLSGCRKS 125  
QY 127 ANSDTHLLQGSULTLTLESPGSPSVQCSPRGKNTIGGKTLVSQLELDGSGTWTCTV 186  
DB 126 PLNS-----VCEWGEPRSTPLTTKA-----VLLVRFQNSPADPQEP 165  
QY 187 LQNKVVEFKIDIVLAFQKASSIVYKKKEGOVEFSPPLAFVTEKLTGSGEL----- 238  
DB 166 QYQGESQKFSQGLAVPEGDSFIVSMCVASSVSGSKSKTQTFQ---GCGILQDPDPANI 222

QY 223 -----W-----MQAERASSSKW--ITPDLNKKEVSVKKTVDPPKLOMGCKLPLAHT 283  
DB 223 TVTAAKRPRLSVTWQDPHSSWSSFYRLRFELKRAERSKTF-----TWMKDLOHCV 278

QY 284 LPQALPOVAGSGLTLAEAKTKLHQEVNLVVMRATOLQKNLTCEWGPSTPKMLSLK 343  
DB 279 IH-----DAMSLRH-----VQLRA---QEEFGQGEWSEWSEAMGTW 315

QY 344 LENK-----EAKYSKREKPPVNLNPAQMWQCLSSGOVLTLESNIKVLPTWSTPVP 399  
DB 316 TESRSPAPANEVS---TPQALITNKDDNLLFRS-----ANATSLPVQ-----AG 360

QY 400 EPKSCDKHTC-----PELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 454  
DB 361 EPKSCDKHTCPCPCAPPELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 420

QY 455 NMVVDGVEVHNAKTKREEQDYNSTYYRVSVLTVLDHQMNGKEVKCKVKSKALPAPIEKT 514  
DB 421 NMVVDGVEVHNAKTKREEQDYNSTYYRVSVLTVLDHQMNGKEVKCKVKSKALPAPIEKT 480

QY 515 ISKAGQPREPQVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 574  
DB 481 ISKAGQPREPQVYTLPPSRDELITKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTP 540

QY 575 PVLDSGSPFLYSKLTVDKSRMQGVNFCSCVMHEALHNHYTQKSLSLSPG 625  
DB 541 PVLDSGSPFLYSKLTVDKSRMQGVNFCSCVMHEALHNHYTQKSLSLSPG 591

## RESULT 56

US-09-499-846-2  
Sequence 2, Application US/09499846  
Patent No. 6656728  
GENERAL INFORMATION:  
APPLICANT: Kavanaugh et al.  
TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR  
TITLE OF INVENTION: RECEPTOR-IMMUNOLOGICAL FUSION  
FILE REFERENCE: 035784/195012 (5784-  
CURRENT APPLICATION NUMBER: US/09/499,846  
CURRENT FILING DATE: 2000-02-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 622  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-499-846-2

Query Match 36.4%; Score 1243; DB 4; Length 622;  
Best Local Similarity 48.3%; Pred. No. 2.2e-90;  
Matches 299; Conservative 47; Mismatches 119; Indels 154; Gaps 23;

QY 109 CEVEDQKEEVOLLVFLGL-TANSDTHLLQGSULTLTLESPGSPSVQCS--RSRGRK---- 161  
DB 55 CRLRDQVQINWLRDQVLAESNRRTITGEVEVQ-DVYPADSGLYACVTSPPSGSDTTY 113  
QY 162 ---NIQGGKTLVSQLELDSDGT-----WTCTVNLQNKVVEFKIDIVYL 202  
DB 114 FSVNVSADALPSSBDDDDDDSSSEKETDNTKRNPAVYWT-----SPEMKKKHAV-- 166  
QY 203 AFQKASSIYKKKEGVEVSFPPLATVE-KLTGSGELMMQAEARASSSKMITFDLKN-XE 260  
DB 167 -----PAAKTIVKRCPSG-----TPNFTLRW-----LNGKE 194  
QY 261 VSVKRVTDPPKLOMG-----KULPLHLTLPQALPOVAGSGLTLALAEKTKLHQEVNL-V 315  
DB 195 FK-----PDHRIGYKRVATWSTIIMDSVP--SDKGVYTCIVENEGSINHVTQLDV 245  
QY 316 VNRATQ--LQKNL-----TCEWGPSTPKMLSLK----- 345  
DB 246 VERSPPRPILQGLPANKTVALGSNVEFMCKVYSDQPHIQMLKHEVNGSKIGPDNLPY 305  
QY 346 ---NKEAKYSKREKPPVNLN-----PEAGMMQCLLSDS----- 375

```
Db 306 VQILKTACVNTDKEMEVLHLRNVSFEDAGEYTCIAGNSIGLSHSAMLTVLEALEERRPA 365
; Sequence 2, Application US/09590656
; Patent No. 6413932
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Borges, Luis G.
; APPLICANT: Fanelow, III, William C.
; TITLE OF INVENTION: TEK ANTAGONISTS
; FILE REFERENCE: 2900-A
; CURRENT APPLICATION NUMBER: US/09/590,656
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/137,889
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-590-656-2

Query Match
Best Local Similarity 81.4%; Score 1237; DB 4; Length 704;
Matches 240; Conservative 9; Mismatches 12; Indels 34; Gaps 5;
```

```
Db 360 VLNPBAGMOCCLSD--SGOVLLESNI--KYLPTWSTVPCPAP----- 399
; Sequence 2, Application US/09590656
; Patent No. 6413932
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Borges, Luis G.
; APPLICANT: Fanelow, III, William C.
; TITLE OF INVENTION: TEK ANTAGONISTS
; FILE REFERENCE: 2900-A
; CURRENT APPLICATION NUMBER: US/09/590,656
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/137,889
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-590-656-2

Query Match
Best Local Similarity 81.4%; Score 1237; DB 4; Length 704;
Matches 240; Conservative 9; Mismatches 12; Indels 34; Gaps 5;
```

```
Db 414 ILPPDSGWCVSNTVAGMVEKPNISVKLPK-----FLNAPVVIDTGHNFVAINISSE 468
; Sequence 2, Application US/09590656
; Patent No. 6413932
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Borges, Luis G.
; APPLICANT: Fanelow, III, William C.
; TITLE OF INVENTION: TEK ANTAGONISTS
; FILE REFERENCE: 2900-A
; CURRENT APPLICATION NUMBER: US/09/590,656
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/137,889
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-590-656-2

Query Match
Best Local Similarity 81.4%; Score 1237; DB 4; Length 704;
Matches 240; Conservative 9; Mismatches 12; Indels 34; Gaps 5;
```

RESULT 58  
US-09-733-764-2

```
; Sequence 2, Application US/09733764
; Patent No. 6521424
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Borges, Luis G.
; APPLICANT: Fanelow, III, William C.
; TITLE OF INVENTION: TEK ANTAGONISTS
; FILE REFERENCE: 2900-A
; CURRENT APPLICATION NUMBER: US/09/733,764
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/590,656
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-733-764-2
```

```
Query Match
Best Local Similarity 81.4%; Score 1237; DB 4; Length 704;
Matches 240; Conservative 9; Mismatches 12; Indels 34; Gaps 5;
```

```
Db 360 VLNPBAGMOCCLSD--SGOVLLESNI--KYLPTWSTVPCPAP----- 399
; Sequence 2, Application US/09499846
; Patent No. 6656728
; GENERAL INFORMATION:
; APPLICANT: Kavanaugh et al.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
; FILE REFERENCE: 035784/195012 (5784-
; CURRENT APPLICATION NUMBER: US/09/499,846
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-499-846-6
```

```
Query Match
Best Local Similarity 66.2%; Score 1235.5; DB 4; Length 497;
Matches 255; Conservative 18; Mismatches 54; Indels 58; Gaps 8;
```

RESULT 59  
US-09-499-846-6



Db	178	GSKIGPDLPRVQILKTAGVTTDKEMEVHLHNRVSPEDAGEYTLDAONISGLSHS---	234
QY	386	VLPTWST-----PVECPAP-----EPKSCDKTHTC-----PELLGSPVF	420
Db	235	---AMLTVLEALERPAVMTSPPLYLESGSGPGLQPKSCDKTHTCPCCPAPELLGSPVF	291
QY	421	LFPKPKDTLMISRPEVTCVAVVDVSHDPEVKFMYVDGVEVHNAKTKPREEQNSTYR	480
Db	292	LFPKPKDTLMISRPEVTCVAVVDVSHDPEVKFMYVDGVEVHNAKTKPREEQNSTYR	351
QY	481	VSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN	540
Db	352	VSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN	411
QY	541	QVSLTCLVKGFPSPDIAEVESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRMOQGN	600
Db	412	QVSLTCLVKGFPSPDIAEVESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRMOQGN	471
QY	601	VFSGSVHREALHNHTOKSLISLSPG 625	
Db	472	VFSGSVHREALHNHTOKSLISLSPG 496	
RESULT 60			
US-09-499-846-4			
Sequence 4, Application US/09499846			
Patent No. 6656728			
GENERAL INFORMATION:			
APPLICANT: Kavanaugh et al.			
TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR			
TITLE OF INVENTION: RECEPTOR-IMMUNOGLOBULIN FUSION			
FILE REFERENCE: 035784/195012 15784-			
CURRENT APPLICATION NUMBER: US/09/499, 846			
CURRENT FILING DATE: 2000-02-07			
NUMBER OF SEQ ID NOS: 12			
SOFTWARE: FastSeq for Windows Version 3.0			
SEQ ID NO 4			
LENGTH: 525			
TYPE: PRT			
ORGANISM: Homo sapiens			
US-09-499-846-4			
Query Match 36.2%; Score 1235.5; DB 4; Length 525;			
Best Local Similarity 66.2%; Pred. No. 6.8e-90;			
Matches 255; Conservative 18; Mismatches 54; Indels 58; Gaps 8;			
QY	293	GSNGTLTLEAKTGLHOC--EVNLVWRATQLOKNL--TCBVGWPTSPKMLSLKLE--	345
Db	146	GSINHTYQLDVVERSPHRPILOAGLRPNKTYALGSNVEFMCKVYSDPOPHIOMLKHIEN	205
QY	346	-----NKEAKVSKREKRVWVLN-----PEAGWQCLSDSGVLLSENIK	385
Db	206	GSKIGPDLPRVQILKTAGVTTDKEMEVHLHNRVSPEDAGEYTLDAONISGLSHS---	262
QY	386	VLPTWST-----PVECPAP-----EPKSCDKTHTC-----PELLGSPVF	420
Db	263	---AMLTVLEALERPAVMTSPPLYLESGSGPGLQPKSCDKTHTCPCCPAPELLGSPVF	319
QY	421	LFPKPKDTLMISRPEVTCVAVVDVSHDPEVKFMYVDGVEVHNAKTKPREEQNSTYR	480
Db	320	LFPKPKDTLMISRPEVTCVAVVDVSHDPEVKFMYVDGVEVHNAKTKPREEQNSTYR	379
QY	481	VSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN	540
Db	380	VSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKN	439
QY	541	QVSLTCLVKGFPSPDIAEVESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRMOQGN	600
Db	440	QVSLTCLVKGFPSPDIAEVESNGQPENNYKTTTPVLDSGSEFLYSKLTVDKSRMOQGN	499
QY	601	VFSGSVHREALHNHTOKSLISLSPG 625	

```

RESULT 61
PCT-US95-03866-12
; Sequence 12, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: Cytomed, Inc. (all states except US)
; APPLICANT: Nocke, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Cytomed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-03866-12

Query Match      36.1%; Score 1232.5; DB 5; Length 424;
Best Local Similarity 88.9%; Pred. No. 8.6e-90;
Matches 232; Conservative 6; Mismatches 18; Indels 5; Gaps 1;

QY 370 CLASDQGVLLSNIKYLPTWSTPVPVPCAPAEPRKSCDXTHTC-----PELLGSPVFLFPP 424
DB 163 CVVSTSLSPENDSRVSVSTKPFMLPVPVADPEPRKSCDXTHTCPCPAPPELLGSPVFLFPP 222
QY 425 KPQKTLMISTRPEYTCVVVDVSHEDPEVKNNWYVDGVEVNNMATTKPREEQYNSTRVVSV 484
DB 223 KPQKTLMISTRPEYTCVVVDVSHEDPEVKNNWYVDGVEVNNMATTKPREEQYNSTRVVSV 282
QY 485 LTVAHQWMLNGKEVCKCVSNKALPAPIEKTISRKAGQPREPOVYTTLPSPRDELTKQVSL 544
DB 283 LTVAHQWMLNGKEVCKCVSNKALPAPIEKTISRKAGQPREPOVYTTLPSPRDELTKQVSL 342
QY 545 TCLVKGFPSPDIAYEWESNGQPENNYKTTTPPVLDSDGSFLLYSKLTVYDKSRMOGQNVFSC 604
DB 343 TCLVKGFPSPDIAYEWESNGQPENNYKTTTPPVLDSDGSFLLYSKLTVYDKSRMOGQNVFSC 402
QY 605 SVMHEALHNHYTKQSLSLSPG 625
DB 403 SVMHEALHNHYTKQSLSLSPG 423

RESULT 62

```

PCT-US95-03866-14  
Sequence 14, Application PC/TUS9503866  
GENERAL INFORMATION:  
APPLICANT: Cytomed, Inc. (all states except US)  
APPLICANT: Nocka, Karl (US only)  
APPLICANT: Lobell, Robert B (US only)  
TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03866  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/220,379  
FILING DATE: 28-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley Jr, James F  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: Cytomed/2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-03866-14

Query Match 36.1%; Score 1232.5; DB 5; Length 424;  
Best Local Similarity 88.9%; Pred. No. 8.6e-90;  
Matches 232; Conservative 6; Mismatches 18; Indels 5; Gaps 1;

QY 370 CLTSDSGQVLLSNIKVLPTWSTVPVPCPAPRPSKCDKTHTC-----PELLGSPSVFLFP 424  
DB 163 CVVSTSLSPKDSRVSYTKPFMLPRVADPEPKSCDKTHTCPCPAPRPSKCDKTHTC-----PELLGSPSVFLFP 222  
QY 425 KPKDTLMSRTPEVTCVVVDVSHEDPEVKFNWYDGVENNAKTRPREQYNSTYRVVSV 484  
DB 223 KPKDTLMSRTPEVTCVVVDVSHEDPEVKFNWYDGVENNAKTRPREQYNSTYRVVSV 282  
QY 485 LTVHOMLNKEKCKKYSNKAALPAETKTSKAKGQPREQVYTLTPSRBELTKNQVSL 544  
DB 283 LTVHOMLNKEKCKKYSNKAALPAETKTSKAKGQPREQVYTLTPSRBELTKNQVSL 342  
QY 545 TCLVKGFPSPDIAVEMESNGQPENNYKTTPEVLDSGDFLYSLKLTVDKSMQGNFVSC 604  
DB 343 TCLVKGFPSPDIAVEMESNGQPENNYKTTPEVLDSGDFLYSLKLTVDKSMQGNFVSC 402  
QY 605 SVMHEALHNHYTKQSLSPG 625  
DB 403 SVMHEALHNHYTKQSLSPG 423

RESULT 63  
US-08-157-101A-7  
Sequence 7, Application US/08157101A  
Patent No. 5808032  
GENERAL INFORMATION:

APPLICANT: KURIHARA, TATSUYA  
APPLICANT: MATSUKURA, SHIGEKAZU  
APPLICANT: TATSUKURA, NOBUO  
APPLICANT: ARIHA, KENJI  
APPLICANT: NISHIHARA, TATSURO  
TITLE OF INVENTION: ANTI-HBs ANTIBODY GENES AND EXPRESSION  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY, MADISON & SUTRO  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/157,101A  
FILING DATE: 05-APR-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: TITUS, MARIANA K  
REGISTRATION NUMBER: 35843  
REFERENCE/DOCKET NUMBER: 9437/204199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3711  
TELEFAX: 202-822-0944  
TELEX: 6714627 CUCH  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 459 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-157-101A-7

Query Match 36.1%; Score 1232.5; DB 1; Length 459;  
Best Local Similarity 46.7%; Pred. No. 9.7e-90;  
Matches 290; Conservative 41; Mismatches 91; Indels 199; Gaps 21;

QY 25 GNRKVLGKKGDTVELTCTAS--QKSIQFM-----KSNQIKIL--GNQGSFL--TK 71  
DB 17 GGGV--QPRSLRLSCAAGFTFSSNSMHWRAQPGKLEWAVILYDGNHKEFYADSVK 74  
QY 72 GPKSLDRADSRSLMDQGNFPLIKLKLTEDSDTYICEVEDQKEVQLVFGLTANSDT 131  
DB 75 GRFTIS--RDNSSKNTLY-----LEVKSLQEDYGVVYC--IRDQ-----TYGV----- 113  
QY 132 HLQ--GQSLTLTLESPPGSSPSVOCSPGKNIQGGKTLVSQLELDGSGTWCTVLQN 189  
DB 114 HRDPSMGQGLTVVSSASTGSPVFLAPSSTSG--TALGCL----- 157  
QY 190 QKVEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFTVEKLTGSGELMWQARASSK 249  
DB 158 -----VKQYFPEPVTVS-----MNSGALASG-- 178  
QY 250 SWITFLDKNKEVSKRVTPDKLQMGKLPALHLLTPALFOYAGSGNLTALAEKTKLH 309  
DB 179 -----VH--TFPAVL--QSSGLYSLSSVTVVPPSSLG 206  
QY 310 QEVNLYVMRATQLOKNULTCEVWGPTSPKMLSLKLEKKAQVSKREKPVVVLNPEAGMWQ 369  
DB 207 TQTYI-----CNV-----NHRKSNTRKVDKV----- 227  
QY 370 CLTSDSGQVLLSNIKVLPTWSTVPVPCPAPRPSKCDKTHTC-----PELLGSPSVFLFP 424  
DB 228 -----EPKSCDKTHTCPCPAPRPSKCDKTHTC-----PELLGSPSVFLFP 257





```

; FRAGMENT TYPE: complete sequence
;
; FEATURE:
;   NAME/KEY: c1CAM(185)/19G fusion protein
;   OTHER INFORMATION: amino acid residues 1-453 =
;   OTHER INFORMATION: c1CAM(453); amino acid residues 454-680 = amino
;   OTHER INFORMATION: acid residues 216-442 of human Ig1 heavy chain
;
US-08-227-496C-15

```

```

Query Match      35.8%; Score 1223.5; DB 3; Length 680;
Best Local Similarity 50.5%; Pred. No. 8.9e-89;
Matches 285; Conservative 44; Mismatches 114; Indels 121; Gaps 18;

```

```

QY 99 LKIEDSDTYICEVEDQKEEVLVFGI--TANSDTHLQGSLLTLESPPSSPSYQCR 156
DB 200 LEVDGDTGVCSID-----GLPVSSEAOVHLALDQRL-----NPTV--- 236
QY 157 SPRGKNIQGGKTLVSQLELDQSGT--WTCTVLQNKQKVEFKIDIVLAFQKASSIVYKK 214
DB 237 -TVGNDSFSAKA-SVS-VTAEDEGTQRLTCVILGNQSGTTLQTVITISFPANVILTKP 293
QY 215 EGEQVEFSPPLAFTVEKLTGSGELMWQAEKASSKMTTDLNKEYSVK-----RTV 267
DB 294 EVSE-----GTEVTYVCEAHPPRAKYT 314
QY 268 QD--PKLQMGKRLPLHLTLPOALPOYAG--SGNLTLEAKTGKLEHVNVLVMMRATOL 332
DB 315 LNVGPAPLPRAQL--LTKATPEBNGSFCSCATLEVAQOLHKQTRLEKLVLYGPRL 371
QY 323 -----QKLTCEVWGPTSPKLMLSLKLENKAKVSKREKPVWVLANPEAGM 367
DB 372 DERDCEGNWTWPEMSQOTPCQAMGNPLPELK-CLNKGTFPLPIG--ESETVTRDLEGT 427
QY 368 WQC-LISDSQVLTLESNIKLPWTSTFVPCPAPBPSSCDKTHTC-----PELLGGPSVFL 421
DB 428 YLCRASTQGEVTRKTVVNVL-----SPRYEDKTHTCPCPAPBELLGGPSVFL 475
QY 422 FPKPKQDTLMISRPETVCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRY 481
DB 476 FPKPKQDTLMISRPETVCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRY 535
QY 482 VSVLTVLHQMMLNGKEKCKVSNKALPAPIEKTISKAKGPRPEOVYTLPPSRDELTKQ 541
DB 536 VSVLTVLHQMMLNGKEKCKVSNKALPAPIEKTISKAKGPRPEOVYTLPPSRDELTKQ 595
QY 542 VSVLTCLVKGFPSPDIAMVESNNGOPENNNTTPPVLDSDGSFPLYSKLTVDKSRMOQGNV 601
DB 596 VSVLTCLVKGFPSPDIAMVESNNGOPENNNTTPPVLDSDGSFPLYSKLTVDKSRMOQGNV 655
QY 602 FSCSVMEALHNHYTOKSLSPG 625
DB 656 FSCSVMEALHNHYTOKSLSPG 679

```

## RESULT 68

```

US-07-934-373C-22
; Sequence 22: Application US/07934373C
; Patent No. 5821337

```

```

; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA

```

```

; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,373C
; FILING DATE: 21-Aug-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-07-934-373C-22

```

```

Query Match      35.8%; Score 1222.5; DB 2; Length 454;
Best Local Similarity 46.8%; Pred. No. 5.9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;

```

```

QY 30 LGKKGDTVELTCTASQKKSIOF--HWKNSQIKILGNQGSFLTK-SPSKLNDRAISRSL 86
DB 11 LVKPGASVAKISCKTSQTYTETTHMMKSHGKSLWEMIGFNPKNQSGSHNQRFMDKATL 70
QY 87 ---WDQGFPLIINKIKIEDSDTYICEVEDQKEEVLVFGITANSDTHLQ--GQSLTL 141
DB 71 AVDKSTAYMELRSLTSEDSGIYIC-----ARWRGLNYGFDVRYFDVWGAGTIV 120
QY 142 TLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSGTWTCTVLQNKQKVEFKIDIV 201
DB 121 TVSSASTKGPVPLAPSSKSTSG-TAALGCL----- 152
QY 202 LAFQKASSIVYKKEGQVEFPLAFTVEKLTGSGELMWQAEKASSKMTTFDLNKEY 261
DB 153 -----VDYFPEPVTVS-----WNSGLTSG----- 173
QY 262 SVKRVTDQPKLQMGKRLPLHLTLPOALPOYAGSGNLTLEAKTGKLEHVNVLVMMRATO 321
DB 174 -----VH-TFPAYL-QSSGLYSLSVTVVPSSSLGTQTYI----- 206
QY 322 LQKNLTCFVWGPTSPKLMLSLKLENKAKVSKREKPVWVLANPEAGMQLSDSGVLE 381
DB 207 -----CNV-----NHKPSNTKVDKAV----- 222
QY 382 SNIKVLPTWSTVPCPAPBPSSCDKTHTC-----PELLGSPSVFLPPPKQDTLMISRTP 436
DB 223 -----EPKSDCKTHTCPCPAPBELLGSPSVFLPPPKQDTLMISRTP 264
QY 437 EYTCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRYVSVLTVLHQMMLNGK 496
DB 265 EYTCVVVDVSHEDPEVKFMWYDGVENNAKTKPREEOYNSTRYVSVLTVLHQMMLNGK 324
QY 497 EYCKVSNKALPAPIEKTISKAKGPRPEOVYTLPPSRDELTKQVSLTCLVKGFPSPDI 556
DB 325 EYCKVSNKALPAPIEKTISKAKGPRPEOVYTLPPSRDELTKQVSLTCLVKGFPSPDI 384
QY 557 AVEMESNNGOPENNNTTPPVLDSDGSFPLYSKLTVDKSRMOQGNVSCSVMEALHNHYT 616
DB 385 AVEMESNNGOPENNNTTPPVLDSDGSFPLYSKLTVDKSRMOQGNVSCSVMEALHNHYT 444
QY 617 QKSLSLSPG 625
DB 445 QKSLSLSPG 453

```

```
RESULT 69
US-08-437-642B-22
/ Sequence 22, Application US/08437642B
/ Patent No. 6054297
/ GENERAL INFORMATION:
/ APPLICANT: Paul J. Carter
/ APPLICANT: Leonard G. Presta
/ TITLE OF INVENTION: Immunoglobulin Variants
/ NUMBER OF SEQUENCES: 47
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 1 DNA Way
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WinPatIn (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/437,642B
/ FILING DATE: 09-May-1995
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/934373
/ FILING DATE: 21-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/146206
/ FILING DATE: 17-NOV-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US92/05126
/ FILING DATE: 15-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/715272
/ FILING DATE: 14-JUN-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Lee, Wendy M.
/ REGISTRATION NUMBER: 40,378
/ REFERENCE/DOCKET NUMBER: P0709P2C1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650/225-1994
/ TELEFAX: 650/952-9881
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 454 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ US-08-437-642B-22

Query Match      35.8%; Score 1222.5; DB 3; Length 454;
Best Local Similarity 46.8%; Pred. No. 5.9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;

QY 30 LGKGGDTVELTCTASQKKSIQF--HMKNNOIKTLNQSGFLTK-GPSKLNDRADSRSL 86
DB 11 LVKPGASVKISCKTSGYTFLEYTHMMKSHGKSLKLEWIGGNPFGSGSSHQRPMDFATL 70
QY 87 ---WDQGNFLLINKLIEDSDTYICEVEDQKEVQLLVGLTANSTHLLQ--GQSLTL 141
DB 71 AUDKSTSTAWELSLTSEDSGIYYC-----ARMKLNFGFVRYFDFWGAQCTTV 120
QY 142 TLSPSPSSPSVQCRSPRGKNIQGGKTLVSQLELDQGTWCTVLQNKVKFEKIDIV 201
DB 121 TVSSASTKGPSVFPLAASSKSTSGG-TAALGCL----- 152
QY 202 LAFOKASSIYYKKEGEOVERSFPLAFVYEKLTSGGELMWOAERASSSKSWITFDLNKEV 261
DB 153 -----VKDYFFPEPVTVS-----MNSGALTSG----- 173
QY 262 SVKAVTQDPKLGWKKLPLHLTLRQALRYVAGSGLTLALEAKTKGLHGEVNLVYMRATQ 321
```

```
DB 174 -----VH-TFPAVL-QSSGLXSLSVTVTPSSSLGTQTYI----- 206
QY 322 LQKNLTCEWGWPTSPKMLSLKLENKAKYSKREKPYVWLNPEAGMKQCLSSGQVLL 381
DB 207 -----CNV-----NHPKSNTKYDKV----- 222
QY 382 SNIKVLPTWSTPVCAPAPBPKSCDKTHTC-----PELLGSPVFLPPPKDITLMSRTP 436
DB 223 -----EPKSCDKHTCCPCAPPELLGSPVFLPPPKDITLMSRTP 264
QY 437 EVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGK 496
DB 265 EVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGK 324
QY 497 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 556
DB 325 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDEMTKNQVSLTCLVKGFYPSDI 384
QY 557 AVEMESNGQPENNYKTTPTVLDSDGSEFLYSKLTVDKSRMQQGNVSCVMHEALHNHYT 616
DB 385 AVEMESNGQPENNYKTTPTVLDSDGSEFLYSKLTVDKSRMQQGNVSCVMHEALHNHYT 444
QY 617 QKSLSLSPG 625
DB 445 QKSLSLSPG 453

RESULT 70
US-08-146-206C-22
/ Sequence 22, Application US/08146206C
/ Patent No. 6407213
/ GENERAL INFORMATION:
/ APPLICANT: Carter, Paul J.
/ APPLICANT: Presta, Leonard G.
/ TITLE OF INVENTION: Method for Making Humanized Antibodies
/ NUMBER OF SEQUENCES: 26
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 1 DNA Way
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: WinPatIn (Genentech)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/146,206C
/ FILING DATE: 17-NO. 6407213-1993
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/715272
/ FILING DATE: 14-JUN-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Lee, Wendy M.
/ REGISTRATION NUMBER: 40,378
/ REFERENCE/DOCKET NUMBER: P0709P1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650/225-1994
/ TELEFAX: 650/952-9881
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 454 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ US-08-146-206C-22

Query Match      35.8%; Score 1222.5; DB 4; Length 454;
Best Local Similarity 46.8%; Pred. No. 5.9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;
```

```

QY 30 LGKKGDTVELTCTAOSQKSIQF--HWKNSNQIKILGNQGSFLTK-GPSKLNDRADSRRL 86
D 11 LVKPGASVKISCKTSGYTFEYTMHMKOSHGLSLMIWGFNPKNGSSSHNQRFMDKATL 70
QY 87 ---WDQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQ--GQSLTL 141
D 71 AVDKSTSTAYMELRSLTSEDSGIYYC-----ARMGGLNYGDFVRFDWAGGATTV 120
QY 142 TLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWCTVLQNKQVEFKIDIYV 201
D 121 TVSSASTGKPSVFLPAPSSKTSGG--TALGCL----- 152
QY 202 LAFQKASSIYKKEGEVERSFPLAFYVEKLTGSGELMWQABRASSSKMITFDLKNKEY 261
D 153 -----VKDYFPEPVTVS-----NMSGALTSG----- 173
QY 262 SVKRYTQDPKLGKMLPLHLTPQALPOYAGSGNLTLLAEKTKGLHGEVNLVWVRATQ 321
D 174 -----VH-TFPAVL-QSSGLYSLSSVTVPSSSLGTQTYI----- 206
QY 322 LQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLNPEAGMWQCLSDSGVLLB 381
D 207 -----CNV-----NHKPSNTKYDKKV----- 222
QY 382 SNIKVLPWTSTVPCPAPPEKSCDKTHTC-----PELLGSPVFLFPKPKDTLMSRTP 436
D 223 -----EPKSCDKTHTCPPCPAPPELLGSPVFLFPKPKDTLMSRTP 264
QY 437 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 496
D 265 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 324
QY 497 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDI 556
D 325 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDI 384
QY 557 AVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDSKRWQGNVSCSVMEHALNHHYT 616
D 385 AVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDSKRWQGNVSCSVMEHALNHHYT 444
QY 617 QKSLSLSPG 625
D 445 QKSLSLSPG 453

```

RESULT 71  
 US-09-705-686-22  
 ; Sequence 22, Application US/09705686  
 ; Patent No. 6639055  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Carter, Paul J.  
 ; Presta, Leonard G.  
 ; TITLE OF INVENTION: Method for Making Humanized Antibodies  
 ; NUMBER OF SEQUENCES: 26  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 1 DNA Way  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: WinPatIn (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/705,686  
 ; FILING DATE: 02-No. 6639055-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/146206

```

; FILING DATE: 17-NOV-1993
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709PID3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-705-686-22

Query Match      35.8%; Score 1222.5; DB 4; Length 454;
Best Local Similarity 46.8%; Pred. No. 5,9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;

QY 30 LGKKGDTVELTCTAOSQKSIQF--HWKNSNQIKILGNQGSFLTK-GPSKLNDRADSRRL 86
D 11 LVKPGASVKISCKTSGYTFEYTMHMKOSHGLSLMIWGFNPKNGSSSHNQRFMDKATL 70
QY 87 ---WDQGNFPLIKNLKIEDSDTYICEVEDQKEEVQLVFGLTANSDFHLLQ--GQSLTL 141
D 71 AVDKSTSTAYMELRSLTSEDSGIYYC-----ARMGGLNYGDFVRFDWAGGATTV 120
QY 142 TLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSGTWCTVLQNKQVEFKIDIYV 201
D 121 TVSSASTGKPSVFLPAPSSKTSGG--TALGCL----- 152
QY 202 LAFQKASSIYKKEGEVERSFPLAFYVEKLTGSGELMWQABRASSSKMITFDLKNKEY 261
D 153 -----VKDYFPEPVTVS-----NMSGALTSG----- 173
QY 262 SVKRYTQDPKLGKMLPLHLTPQALPOYAGSGNLTLLAEKTKGLHGEVNLVWVRATQ 321
D 174 -----VH-TFPAVL-QSSGLYSLSSVTVPSSSLGTQTYI----- 206
QY 322 LQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLNPEAGMWQCLSDSGVLLB 381
D 207 -----CNV-----NHKPSNTKYDKKV----- 222
QY 382 SNIKVLPWTSTVPCPAPPEKSCDKTHTC-----PELLGSPVFLFPKPKDTLMSRTP 436
D 223 -----EPKSCDKTHTCPPCPAPPELLGSPVFLFPKPKDTLMSRTP 264
QY 437 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 496
D 265 EYTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQVNSTYRVVSVLTVLHQDMLNGK 324
QY 497 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDI 556
D 325 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPSRDELTKNQVSLTCLVKGFYPSDI 384
QY 557 AVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDSKRWQGNVSCSVMEHALNHHYT 616
D 385 AVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDSKRWQGNVSCSVMEHALNHHYT 444
QY 617 QKSLSLSPG 625
D 445 QKSLSLSPG 453

```

RESULT 72  
 PCT-US93-07832-22  
 ; Sequence 22, Application PC/TUS9307832  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Genentech, Inc.  
 ; TITLE OF INVENTION: Immunoglobulin Variants



```

: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/07832
: FILING DATE: 19930820
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/715272
: FILING DATE: 14-JUN-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/05126
: FILING DATE: 15-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/934373
: FILING DATE: 21-AUG-1992
: ATTORNEY/AGENT INFORMATION:
: NAME:
: REGISTRATION NUMBER:
: REFERENCE/DOCKET NUMBER: 709P2PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE:
: TELEFAX: 415/952-9881
: TELEEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 454 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: FCT-US93-07832-22

Query Match      35.8%; Score 1222.5; DB 5; Length 454;
Best Local Similarity 46.8%; Pred. No. 5.9e-89;
Matches 285; Conservative 35; Mismatches 110; Indels 179; Gaps 15;

QY 30 LGKGGTVLCTCTASQKSIQF--HWKNSNOIKILGNGSFLTK-GPSKINDRADSRSL 86
DB 11 LVKGGASVYKISCTSGYFTFTYTHMMKQSHGKSLKEMIGFNPKNKGSSHNQRFMDKATL 70
QY 87 ---WDQGNFPLIKLKIEDSDTYICEVEDQKEVQLLVFGLTANSPTHLQ--GQSILTL 141
DB 71 AVDKSTSTAVYMLNSLTSDESGIYYC-----ARWRGLNYGFDRYFDVWGAGTTV 120
QY 142 TLSEPPGSSPVQCCSPRGNIOGGKTLVSQLELDGSGTWTCTVLQNKVFEKIDIV 201
DB 121 TVSSASTGSPVFLAPSSKSTSGG-TAALGL----- 152
QY 202 LAFQASSIVYKKEGQVEFSFPLAFVTEKLTSGELMWAERASSKSWITFDLKKEV 261
DB 153 ---VKDYFPEPVTVS-----WNSGALTSG----- 173
QY 262 SVKAVTQDPKIQMGKKLPLHLITLQALFQYAGSGNLTALEAKGKLGHOENVLVMMRATQ 321
DB 174 ---VH--TFPAVL--QSSGLYSLSSVTVVPSSSLGTQTYI----- 206
QY 322 LQKNLTCEVMGPTSPKMLSLKLENKEAKVSKREKPVVVLNPEAGMOCCLSDSGQVLE 381
DB 207 ---CNV-----NHKSNTKVDKKV----- 222
QY 382 SNIKVLPWSTPVPCPAPEPKSCDKTHTC-----PELLGSPVFLFPKPKDTIMISRT 436
DB 223 ---EPKSCDKTHTCPCPAPELLGSPVFLFPKPKDTIMISRT 264

```

```

QY 437 EVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQYNSTYRVSVLTFLVHQDWLNGK 496
DB 265 EVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTRPREQYNSTYRVSVLTFLVHQDWLNGK 324
QY 497 EYKCKSNKKLPAPIEKTTISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 556
DB 325 EYKCKSNKKLPAPIEKTTISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDI 384
QY 557 AVEMESNGOPENNYKTPPLTLDSDGSPFLYSKLTVDKSRMQQGVFSCVWHEALHNHYT 616
DB 385 AVEMESNGOPENNYKTPPLTLDSDGSPFLYSKLTVDKSRMQQGVFSCVWHEALHNHYT 444
QY 617 QKSLSLSPG 625
DB 445 QKSLSLSPG 453

RESULT 73
US-09-740-002-25
: Sequence 25, Application US/09740002
: Patent No. 6537809
: GENERAL INFORMATION:
: APPLICANT: BRAMS, PETER
: TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES
: TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
: TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF
: FILE REFERENCE: 037003-0275759
: CURRENT APPLICATION NUMBER: US/09/740,002
: CURRENT FILING DATE: 2000-12-20
: PRIOR APPLICATION NUMBER: 09/335,697
: PRIOR FILING DATE: 1999-06-18
: PRIOR APPLICATION NUMBER: 08/488,376
: PRIOR FILING DATE: 1995-06-07
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn ver. 2.1
: SEQ ID NO 25
: LENGTH: 475
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-740-002-25

Query Match      35.8%; Score 1222.5; DB 4; Length 475;
Best Local Similarity 45.6%; Pred. No. 6.4e-89;
Matches 293; Conservative 35; Mismatches 111; Indels 203; Gaps 17;

QY 10 LLVLTQLALPAAATQGNKVVLGKKGDTVELCTAS-----QKSIQPHWK 54
DB 10 LVNAATRVLSQVQJQESGPPVVVKTETLTCTVSGFSLNPRMGVTWIRQPEKALEW 68
QY 55 NSNOIKILGN-----QGSFLTKGPSKLNDRADSRSLMDQGNFPLIKLKIEDSDTYIC 109
DB 69 -----LGNIFSSDEKSFSPSLKSRLLTTSQDTSRS-----QVLSLTNVDPVDTATYYC 116
QY 110 EVEDQKEVQLLVGLTANSPTHL-LOGQSILTLTLESPPSSPVQCCSPRGNIOGGK 168
DB 117 -----ARGLVDINAYLYLYDYGQGLTVVSSASTKGSVPFLAPSSKSTSGG-T 167
QY 169 LSVSQLELDGSGTWTCTVLQNKVFEKIDIVLAFQASSIVYKKEGQVEFSFPLAFT 228
DB 168 AALGCL-----VKDYFPEPVTVS----- 183
QY 229 VEKLTSGGELMWAERASSKSWITFDLKKEVSVKAVTQDPKIQMGKKLPLHLITLQAL 288
DB 184 VS-----WNSGALTSG-----VH--TFPAVL 202
QY 289 PQVAGSNLTALAEAKGKLGHOENVLVMMRATQIQKNLTCEVMGPTSPKMLSLKLENKE 348
DB 203 ---QSSGLYSLSSVTVVPSSSLGTQTYI-----CNV----- 230
QY 349 AKVSKREKPVVVLNPEAGMOCCLSDSGQVLESNIKVLPWSTPVPCPAPEPKSCDKTH 408
DB 231 ---NHKPF-----SNTKV-----DKAPEPKSCDKTH 252

```

QY 409 TC-----PELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEV 463  
DB 253 TCPCCAPPELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEV 312  
QY 464 HNAKTKREBOVNSTYRVSVLTVLHODMNLGKEVKCKVSKNKAAPAEIKTISKAKGQPR 523  
DB 313 HNAKTKREBOVNSTYRVSVLTVLHODMNLGKEVKCKVSKNKAAPAEIKTISKAKGQPR 372  
QY 524 EPQVYITLPPSRDELTKQVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGSF 583  
DB 373 EPQVYITLPPSRDELTKQVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGSF 432  
QY 584 FLYSKLTVDKSRMOQGNVFSQSVHMEALHNHYTQKSLSLSPG 625  
DB 433 FLYSKLTVDKSRMOQGNVFSQSVHMEALHNHYTQKSLSLSPG 474

RESULT 74  
US-09-049-672A-4  
Sequence 4, Application US/09049672A  
Patent No. 6135941  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Tang, Y. Tom  
APPLICANT: Yue, Henry  
APPLICANT: Au-Yang, Janice  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/049,672A  
FILING DATE: HEREWITH  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Cerrone, Michael C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0497 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 473 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PANTUT01  
CLONE: 1513264  
US-09-049-672A-4

Query March 35.8%; Score 1221.5; DB 3; Length 473;  
Best Local Similarity 45.3%; Pred. No. 7.6e-89;

Matches 288; Conservative 40; Mismatches 125; Indels 183; Gaps 15;  
QY 8 RHLLLVQLALP-----AATQGNRVVLGKGDYVELTCTAS--QKSIQFHKRNSQI 59  
DB 2 KHLMEFLLVAAPRWYLSQVQLQESGPGLVKSEETLSLCAVSGSITSGGYWMIWRP 61  
QY 60 KILGNO--GSFLTKGPSKNDRAISRSL---WDQGNPLIITKNLKIETSDTYICVEDQ 114  
DB 62 PGKGLMEITGIYISGSLYNPSLSKSRVITSVDTSKQFSLKLSVTAATATAYYCARDP 120  
QY 115 KEEVOLLVEGLTANSDTHLLQGSITLTLLESPGPSVQCRSPRKNIQGGKTLVSQI 174  
DB 121 -----VGLRGANGMDVWGQGLTVTSASTKSPFPLAPSSKSTSG--TALGCI 171  
QY 175 ELQDSGTWTCYVLQNGKVEFKIDIVLAFOKASSIVYKKEGEQVEFSFPLATVYKLG 234  
DB 172 -----VQDFPEEPTVS----- 183  
QY 235 SGEIWMQAEERASSKSWITFDLKNKEVSRYTDPKLGKGLPLHLTLPOLPYAGS 294  
DB 184 -----INSGALISG-----VH-TFPAVL-QSSGL 205  
QY 295 GNLTALAEAKTGKLGQEVNLVVMRATQLOKNLTCEVWGPTSPKMLSLKLNKEAKVSKR 354  
DB 206 YSLSSVTVTPSSSLGTYI-----CNY-----NHKPSNTKY 237  
QY 355 EKPVWVLANPEAGMOCCLSDSGVLIESNIKVLPTWSTVPCPAPPKSCDKHTTC---- 410  
DB 238 DKRV-----EPKSCDKHTTTPCP 256  
QY 411 -PELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEVHNAKTK 469  
DB 257 APELLGGPSVFLPPPKDITLMISRTPEVTCVVVDVSHEDPEVKFNRYVDGVEVHNAKTK 316  
QY 470 PREQYNSTYRVSVLTVLHODMNLGKEVKCKVSKNKAAPAEIKTISKAKGQPREQVYT 529  
DB 317 PREQYNSTYRVSVLTVLHODMNLGKEVKCKVSKNKAAPAEIKTISKAKGQPREQVYT 376  
QY 530 LPPSRDELTKQVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGSFLYSKL 589  
DB 377 LPPSRDELTKQVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTTTPVLDSDGSFLYSKL 436  
QY 590 TVDKSRMOQGNVFSQSVHMEALHNHYTQKSLSLSPG 625  
DB 437 TVDKSRMOQGNVFSQSVHMEALHNHYTQKSLSLSPG 472

RESULT 75  
US-08-470-299-4  
Sequence 4, Application US/08470299  
Patent No. 5783181  
GENERAL INFORMATION:  
APPLICANT: Browne, Michael J.  
APPLICANT: Murphy, Kay E.  
APPLICANT: Chapman, Conrad G.  
APPLICANT: Clinkenbeard, Helen E.  
APPLICANT: Young, Peter R.  
APPLICANT: Shatzman, Allan R.  
TITLE OF INVENTION: No. 5783181el Compounds  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road, P.O. Box 1539  
CITY: King of Prussia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/470.299
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sutton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE/DOCKET NUMBER: P31005C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5024
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-470-299-4

Query Match      35.7%; Score 1220.5; DB 1; Length 387;
Best Local Similarity 83.4%; Pred. No. 6.8e-89;
Matches 236; Conservative 13; Mismatches 23; Indels 11; Gaps 4;

OY 351 VSKKEKPVWVLPNPEAKMQCLISDSGQVLLSNNIKVLPF--WSTFVPCPA-PEPKSCDKT 407
DB 107 LKRLDRNLMLGL---AGLNSCPVKEANQSTLENFLERLKTIMREKDKSCSGTEPKSADKT 163
OY 408 HTCC-----PELLGSPVFLFPKPKDITMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVE 462
DB 164 HTCPPCPAPELGGSPVFLFPKPKDITMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVE 223
OY 463 VHNATKPREEOYNSTYRVSVSLTVLHODMNLNGKEKCKVSNKALPAPIEKTISKAKQP 522
DB 224 VHNATKPREEOYNSTYRVSVSLTVLHODMNLNGKEKCKVSNKALPAPIEKTISKAKQP 283
OY 523 REPQYTYLPSPRDELTKQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTTPVLDSGS 582
DB 284 REPQYTYLPSPRDELTKQVSLTCLVKGFYPSDIAVEMESNQGPENNYKTTTPVLDSGS 343
OY 583 FFLYSKLTVDKSRMOQGVFSGSVNHEALHNHYTQKSLSLSPG 625
DB 344 FFLYSKLTVDKSRMOQGVFSGSVNHEALHNHYTQKSLSLSPG 386

RESULT 76
US-09-532-856-6
; Sequence 6, Application US/09532856
; Patent No. 6458350
; GENERAL INFORMATION:
; APPLICANT: COSMAN, David J.
; APPLICANT: MULBERG, Jurgen H.
; APPLICANT: FANSLAW III, William C.
; APPLICANT: KUBIN, Marek
; TITLE OF INVENTION: ULBP DNA AND POLYPEPTIDES
; FILE REFERENCE: 2866-USA
; CURRENT APPLICATION NUMBER: US/09/532, 856
; EARLIER FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: PCT/US98/27048
; EARLIER FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/069, 857
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 60/092, 946
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(223)
; OTHER INFORMATION: ULBP-2 sequences
; FEATURE:
```

```

; NAME/KEY: PEPTIDE
; LOCATION: (224)..(453)
; OTHER INFORMATION: Human Ig Fc sequences
US-09-532-856-6

Query Match      35.7%; Score 1220.5; DB 4; Length 453;
Best Local Similarity 58.9%; Pred. No. 8.6e-89;
Matches 258; Conservative 31; Mismatches 70; Indels 79; Gaps 11;

OY 236 GELWQAEKRRASSKSMITFDLAKKEVSKRTYQPKLQMGKKLPLHLTPALQRYAGSG 295
DB 46 GPRWCAVQGVDEKTFPHYDCGNTVT-----FVSPLGKTL----- 81
OY 296 NLTLAEAKTKLHQNVLVWVRATOLQKNLTCVWGPSTP-KLMLSLKENKAKVSK- 353
DB 82 NVTAMKAKQNPVLEKREVDIL---TEQLRDILQLENYTPKEBLTLQARNSCQKAEHSG 137
OY 354 -----REKPVW-VLNPEA-----GMWQCL--LSDS 375
DB 138 SWGSPDQQLFLFDSEKRWMTTVHPGARKKKEKENDKVAMS FHYFSMGDCIGWLED- 196
OY 376 GQVLLSNNIKVLPWSTFVPCPAEP--KSCDKHTTC-----PELLGSPVFLFPKPK 427
DB 197 --FLMGMDSTLEPESAGAPLAMS SGTTLRRSCDKHTTCPCPAPEAEAGAPSVFLFPKPK 254
OY 428 DTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNATKPREEOYNSTYRVSVSLTV 487
DB 255 DTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNATKPREEOYNSTYRVSVSLTV 314
OY 488 LHDMLNGKEKCKVSNKALPAPIEKTISKAKQPREQYTYLPSPRDELTKQVSLTCL 547
DB 315 LHDMLNGKEKCKVSNKALPAPIEKTISKAKQPREQYTYLPSPRDELTKQVSLTCL 374
OY 548 VKGFYPSDIAVEMESNQGPENNYKTTTPVLDSGSFFLYSKLTVDKSRMOQGVFSGSV 607
DB 375 VKGFYPSDIAVEMESNQGPENNYKTTTPVLDSGSFFLYSKLTVDKSRMOQGVFSGSV 434
OY 608 HEALHNHYTQKSLSLSPG 625
DB 435 HEALHNHYTQKSLSLSPG 452

RESULT 77
US-09-524-100C-6
; Sequence 6, Application US/09524100C
; Patent No. 6653447
; GENERAL INFORMATION:
; APPLICANT: COSMAN, David J.
; APPLICANT: MULBERG, Jurgen H.
; APPLICANT: FANSLAW III, William C.
; APPLICANT: KUBIN, Marek
; APPLICANT: ARMSTRONG, Richard J.
; TITLE OF INVENTION: ULBP DNA AND POLYPEPTIDES
; FILE REFERENCE: 2866-US
; CURRENT APPLICATION NUMBER: US/09/524, 100C
; EARLIER FILING DATE: 2002-05-21
; EARLIER APPLICATION NUMBER: PCT/US98/27048
; EARLIER FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: US 60/069, 857
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: US 60/092, 946
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(223)
```

OTHER INFORMATION: ULBP-2 sequences  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (224) ..(453)  
OTHER INFORMATION: Human Ig Fc sequences  
US-09-524-100C-6

Query Match 35.7%; Score 1220.5; DB 4; Length 453;  
Best Local Similarity 58.9%; Pred. No. 8.6e-89;  
Matches 258; Conservative 31; Mismatches 70; Indels 79; Gaps 11;

236 GELMWAERASSSKSWITFLKNKEVSKRVTPDKLQMGKLLPLHLTPQALPQYAGSG 295  
46 GPRMCAVQGVDEKTFPHYDCGKTYT-----PVSPLQKL----- 81  
296 NLTLALAKTGKLEHVNLYVMRATOLQKNLTCEVWGPTSP-KLMLSLKLENKAVSK- 353  
82 NNTTAMKAKNPVIREVVDIL-----TEQLRDIOLENTYPRKEPLTLQARMSCQKAEHSG 137  
354 -----RKKPW-VLNPA-----GMMOCL--LSDS 375  
138 SWQSPFDGQIFLLPDSSEKRWTTVHPGARCKMEKENDKVAMS FHYFSMGDCIGMLED- 196  
376 GOVLLSENIKVLPTWSTPVPAPAP-----KSCDKTHTC-----PELLGSPVFLPPKPK 427  
197 --FLMGMDSTLEBSAGAPLAMSSTTOLRSCDKHTCPPCAPAEAGASVFLFPKPK 254  
428 DTLMISTPEVTCVVDVSHEDPEVKENMYVDGVEVNAKTKPREQYNSTYRVSVLT 487  
255 DTLMISTPEVTCVVDVSHEDPEVKENMYVDGVEVNAKTKPREQYNSTYRVSVLT 314  
488 LHQDWLNGKRYKCKVSNKALPAPLEKTSKAKQPREPOVYTLPPSRDELTKNOVSLTCL 547  
315 LHQDWLNGKRYKCKVSNKALPAPLEKTSKAKQPREPOVYTLPPSRDELTKNOVSLTCL 374  
548 VKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGVFSGSV 607  
375 VKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGVFSGSV 434  
608 HEALHNHYTKSLSLSPG 625  
435 HEALHNHYTKSLSLSPG 452

RESULT 78  
US-09-499-846-10  
Sequence 10, Application US/09499846  
Patent No. 6656728  
GENERAL INFORMATION:  
APPLICANT: Kavanaugh et al.  
TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR  
FILE REFERENCE: 035784/195012 (5784-  
CURRENT APPLICATION NUMBER: US/09/499, 846  
CURRENT FILING DATE: 2000-02-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 497  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-499-846-10

Query Match 35.7%; Score 1220.5; DB 4; Length 497;  
Best Local Similarity 65.7%; Pred. No. 9.8e-89;  
Matches 253; Conservative 18; Mismatches 56; Indels 58; Gaps 8;

293 GSGNLTALAEAKTGKLEHVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLE-- 345  
118 GGINHTYQLDIVERSPHRPILQAGLPANKTVALGSNVEFMCKVYSDPOPHIOMLKHEVN 177  
346 -----NKEAVSKREKPVWVNL-----PEAGMMOCLLSDSGOVLLESNIK 385

178 GSKIGPNDLPYVQILKTAGVNTTDKEMEVLHLRNVSPEDAGEYTCLAGNSIGLSHHS--- 234  
386 VLPTST-----PVPCAP-----EPKSCDKHTHC-----PELLGSPVFL 420  
235 ---AMLTVLEALEBRPAWVTSPLYLEGSGSPGLQEPKSCDKHTCPPCAPAELEGSPVFL 291  
421 LPPPKRDTLMSRTPEVTCVVDVSHEDPEVKENMYVDGVEVNAKTKPREQYNSTYR 480  
292 LPPPKRDTLMSRTPEVTCVVDVSHEDPEVKENMYVDGVEVNAKTKPREQYNSTYR 351  
481 VVSVTLVTHQDWLNGKRYKCKVSNKALPAPLEKTSKAKQPREPOVYTLPPSRDELTKN 540  
352 VVSVTLVTHQDWLNGKRYKCKVSNKALPAPLEKTSKAKQPREPOVYTLPPSRDELTKN 411  
541 QVSLTCLVYGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 600  
412 QVSLTCLVYGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 471  
601 VFSCSVMEALHNHYTKSLSLSPG 625  
472 VFSCSVMEALHNHYTKSLSLSPG 496

RESULT 79  
US-09-499-846-8  
Sequence 8, Application US/09499846  
Patent No. 6656728  
GENERAL INFORMATION:  
APPLICANT: Kavanaugh et al.  
TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR  
FILE REFERENCE: 035784/195012 (5784-  
CURRENT APPLICATION NUMBER: US/09/499, 846  
CURRENT FILING DATE: 2000-02-07  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 525  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-499-846-8

Query Match 35.7%; Score 1220.5; DB 4; Length 525;  
Best Local Similarity 65.7%; Pred. No. 1.1e-88;  
Matches 253; Conservative 18; Mismatches 56; Indels 58; Gaps 8;

293 GSGNLTALAEAKTGKLEHVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLE-- 345  
146 GGINHTYQLDIVERSPHRPILQAGLPANKTVALGSNVEFMCKVYSDPOPHIOMLKHEVN 205  
346 -----NKEAVSKREKPVWVNL-----PEAGMMOCLLSDSGOVLLESNIK 385  
206 GSKIGPNDLPYVQILKTAGVNTTDKEMEVLHLRNVSPEDAGEYTCLAGNSIGLSHHS--- 262  
386 VLPTST-----PVPCAP-----EPKSCDKHTHC-----PELLGSPVFL 420  
263 ---AMLTVLEALEBRPAWVTSPLYLEGSGSPGLQEPKSCDKHTCPPCAPAELEGSPVFL 319  
421 LPPPKRDTLMSRTPEVTCVVDVSHEDPEVKENMYVDGVEVNAKTKPREQYNSTYR 480  
320 LPPPKRDTLMSRTPEVTCVVDVSHEDPEVKENMYVDGVEVNAKTKPREQYNSTYR 379  
481 VVSVTLVTHQDWLNGKRYKCKVSNKALPAPLEKTSKAKQPREPOVYTLPPSRDELTKN 540  
380 VVSVTLVTHQDWLNGKRYKCKVSNKALPAPLEKTSKAKQPREPOVYTLPPSRDELTKN 439  
541 QVSLTCLVYGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 600  
440 QVSLTCLVYGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 439  
601 VFSCSVMEALHNHYTKSLSLSPG 625  
500 VFSCSVMEALHNHYTKSLSLSPG 524



```

; TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF
; FILE REFERENCE: 037003-0275759
; CURRENT APPLICATION NUMBER: US/09/740.002
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/335,697
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 08/488,376
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 475
; TYPE: PRF
; ORGANISM: Homo sapiens
; US-09-740-002-27

Query Match      35.7%; Score 1219.5; DB 4; Length 475;
Best Local Similarity 45.0%; Pred. No. 1.1e-88;
Matches 290; Conservative 31; Mismatches 116; Indels 207; Gaps 16;

QY 10 LLLVQLALLPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNOIKILGNOGSEL 69
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 10 LVAVATRVLSQVOLQESGPALVVKPTQTLTLCTFS-----GFSLSTRGMSVNM 58
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 70 TKGPSKLAND---RADSRSLMDQNF-----PLIKKLIKIDSDT 106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 59 RQPGKALEWLARD-----WDDDTFYSASLKTRLSKTSKKNQVLRMTNVDPVDTAT 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 107 YICEVEDQKEVOLLVLEGLTANSPDTHLQOSLTLLLESPPGSSPSVOCSPRGNIOGG 166
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 114 YFCARASLYSDSFYLF-----YHAYWQGTIVTVSSASTGSPVFLPASPSSKTSGG 166
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 167 KTLVSQLELDQSGTGTCTVLQNOKKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLA 226
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 167 -TALGCL-----WNSGALTSG-----VKDYFPPBP 181
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 227 FTYEKLTSGBELMWQAEARASSSKSWITFDLKNKEVSXKRTQDPKQLQMGKLLPLHLTL 286
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 182 VTVS-----WNSGALTSG-----VH-TFPA 200
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 287 ALPOYAGSGLTLALAEKTKLHGEVNLVVMRATQLOKNLTCEVWGFTSPKMLSLKEN 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 201 VL-QSSGLYSLSSVTVYVSSSLGTQYI-----CNV----- 230
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 347 KEAKVSKREKPVWVLANPEAGMOCCLSDSGVLIESNIKVLPTWSTVPFPAPEPKSCDK 406
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 231 -----NHRP-----SNWKV-----DKKAEPKSCDK 250
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 407 THTC-----PELIGGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGV 461
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 251 THTCPPCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGV 310
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 462 EVHNAKTPREEOYNSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKG 521
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 311 EVHNAKTPREEOYNSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKG 370
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 522 PREPOYVTLTPPSRDELTKNOVSLTCLVKGYSPSDIAVEMSNQGPENNYKTTTPVLDSG 561
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 371 PREPOYVTLTPPSRDELTKNOVSLTCLVKGYSPSDIAVEMSNQGPENNYKTTTPVLDSG 430
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 582 SFFLYSKLTVDKSRMOCGNVPSGVMEALHNHYTQKSLSLSPG 625
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 431 SFFLYSKLTVDKSRMOCGNVPSGVMEALHNHYTQKSLSLSPG 474
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 83
US-08-487-550-12
; Sequence 12, Application US/08487550
; Patent No. 6113898
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,

```

```

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; TITLE OF INVENTION: IMMUNOSUPPRESSANTS"
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,550
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-487-550-12

Query Match      35.7%; Score 1219.5; DB 3; Length 476;
Best Local Similarity 53.7%; Pred. No. 1.1e-88;
Matches 277; Conservative 31; Mismatches 89; Indels 119; Gaps 15;

QY 171 VSGLELDGSG-----TWTCYVLQNOKKVEFKIDIVLAFOKASSIVYKKEGQVE 220
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 18 LQOVOLQESGPGLVKSSETLTLCAV-----SGGSIS 49
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 221 FSPFLAFTYKLTGSGELMWQAEARASSSKSWITFDLKNKEVSXKRTQDPKQLQMGKLL 280
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 50 GGYGMW-IRQPPGKLEWIGSFYSSSGNTYVNSPKS-QVITIS--TDSKNOFSLKL-- 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 281 HLTLPQALPOYAGSGLTLALBA---KTGKLHGEVNLVVMRATQLOKNLTCEVWGCP--- 333
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 104 -----NSMTAADTAAYVYCVARDLFSVVGMY-----NNWFVWPGVL 141
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 334 -----TSPKMLSLKENKEAKVSKR-----EKPVVY-----LNPEAGMOC 370
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 142 VTVSSASTGSPVFLPAPSSKSTSGTALGCLVMDYFPEPVTVSWSGALTSGVTFPA 201
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 371 LLSDSGQVLLES-----NIKVLPTWSTVPFPAPEPKSCDKTHTC----- 410
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 202 VLQSSGLYSLSSVTVYVSSSLGTQYICVNAHKPS-NTKVDKKA-EPKSCDKTHTCPPCP 259
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 411 -BELIGGSPVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTK 469
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 260 APELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGVVHNAKTK 319
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 470 PREEOYNSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPOYV 529
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 320 PREEOYNSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPOYV 379
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 530 LPPSRDELTKNOVSLTCLVKGYSPSDIAVEMSNQGPENNYKTTTPVLDSGSGFFLYSKL 589
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 380 LPPSRDELTKNOVSLTCLVKGYSPSDIAVEMSNQGPENNYKTTTPVLDSGSGFFLYSKL 439
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 590 TVDKSRMOCGNVPSGVMEALHNHYTQKSLSLSPG 625
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 440 TVDKSRMOCGNVPSGVMEALHNHYTQKSLSLSPG 475
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 84  
US-09-526-098-12  
Sequence 12, Application US/09526098  
Patent No. 6492134  
GENERAL INFORMATION:  
APPLICANT: Anderson, Darrell R.  
TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC  
TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS  
TITLE OF INVENTION: IMMUNOSUPPRESSANTS"  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/526,098  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/383,916  
FILING DATE:  
APPLICATION NUMBER: US 08/487,550  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-131  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 476 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-526-098-12

Query Match 35.7%; Score 1219.5; DB 4; Length 476;  
Best Local Similarity 53.7%; Pred. No. 1.1e-88;  
Matches 277; Conservative 31; Mismatches 89; Indels 119; Gaps 15;

171 VSQELQDSG-----TWCTVLQNKVKFEKIDIVLAFQKASSIVYKKEGEQVE 220  
18 LSGVQLQSGSGGLVPSFETLSLTCAV-----SGGSIS 49  
221 FSPFLATVEKLTSGSGELMWQAEARASSKSWITPDLKXKEYSVKRVTDPLQNGKXLP 280  
50 GGYGNGW-IRQPPGSGEMISFYSSSGNTYNSPLKS-QVTIS--TDSKNGPSLKL- 103  
281 HLTLPLQALPOYAGSGNLTFLAE---KTGKLHQEVLVWMARATQLOKNLCEVWGP--- 333  
104 -----NSMTAADTAVYYCVRDLFSVGVGVY-----NNMFVWGPGL 141  
334 -----TSPKLMSTLKENKEAKVSKR-----EKPVWY-----LNPEAGMWQC 370  
142 VTSSASATKGPSVFLAPSSKSTSGTAALGLVKDYPEPEVYTSWMSGALTSGVHFFPA 201  
371 LLSBSGQVLES-----NIKVLPTWSTPVPCEAPPEKSCDKTHTC----- 410  
202 VLSSSGVLSLSVTVTPSSSLGTOTYICNVNKKPS-NTKVDKKA-EPKSCDKTHTCPCP 259

411 -PELLGSPVLEFPPEKDTLMTSRPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTK 469  
260 APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTK 319  
470 PREQYNSTRYVSVLVTHLDWLNKGEYKCKVSNKALPAPIETKISKAKQPREPVYTT 529  
320 PREQYNSTRYVSVLVTHLDWLNKGEYKCKVSNKALPAPIETKISKAKQPREPVYTT 379  
530 LPPSRDELTKNQVSLTCLVGFPSDIAVEMESNGOEENNYKTTPPVLDSDGSFFLYSKL 589  
380 LPPSRDELTKNQVSLTCLVGFPSDIAVEMESNGOEENNYKTTPPVLDSDGSFFLYSKL 439  
590 TVDKSRWQOGVNFSCSVMEALNHNHYTQKSLSLSPG 625  
440 TVDKSRWQOGVNFSCSVMEALNHNHYTQKSLSLSPG 475

RESULT 85  
US-08-472-888A-7  
Sequence 7, Application US/08472888A  
Patent No. 6613746  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Walz, Gerd  
TITLE OF INVENTION: AGP-ANTIBODY FUSION PROTEINS  
TITLE OF INVENTION: AND RELATED MOLECULES AND METHODS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: Pasteset for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,888A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/618,314  
FILING DATE: 23-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/258001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 442 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-472-888A-7

Query Match 35.6%; Score 1216; DB 4; Length 442;  
Best Local Similarity 89.6%; Pred. No. 1.9e-88;  
Matches 233; Conservative 3; Mismatches 14; Indels 10; Gaps 2;

371 LLSBSGQVLESNIKVLPTWSTPVPCEAPPEKSCDKTHTC-----PELLGSPVLEFPPEK 425  
187 VLSSSGVLSLSVTVTPSSSDKKV-----EPKSCDKTHTCPCPAPPELLGGPSVFLFPPK 241  
426 PKDTLMTSRPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREQYNSTRYVSVLV 485



Db 242 PKDTLMISRTPEVTCVVDVSHEDPEYKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVL 301  
QY 486 TVLHODWLNGKEYKCKVSNKALPAPLEKTKSKAKGQPREPQVYTLPPSRDELTKNQVSLT 545  
Db 302 TVLHODWLNGKEYKCKVSNKALPAPLEKTKSKAKGQPREPQVYTLPPSRDELTKNQVSLT 361  
QY 546 CLVKGFPYSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFSGS 605  
Db 362 CLVKGFPYSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFSGS 421  
QY 606 VMHEALHNHYTKSLSPG 625  
Db 422 VMHEALHNHYTKSLSPG 441

RESULT 86  
PCT-US96-10043-9  
Sequence 9, Application PC/TUS9610043  
GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation  
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES  
TITLE OF INVENTION: AND METHODS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/10043  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000,213  
FILING DATE: 14-JUN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lech, Karen F.  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 00786/284001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 442 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-10043-9

Query Match 35.6%; Score 1216; DB 5; Length 442;  
Best Local Similarity 89.6%; Pred. No. 1.9e-88;  
Matches 233; Conservative 3; Mismatches 14; Indels 10; Gaps 2;

QY 371 LLSDSGQVLLSNKVLPTWSTPVPAPAPKSCDKHTHC-----PELLGSPSVLFPKP 425  
Db 187 VLQSSGSLYSIVTVVSSSDKKV-----EPKSCDKHTTCPPAPAPLLGSPSVLFPKP 241  
QY 426 PKDTLMISRTPEVTCVVDVSHEDPEYKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVL 485  
Db 242 PKDTLMISRTPEVTCVVDVSHEDPEYKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVL 301  
QY 486 TVLHODWLNGKEYKCKVSNKALPAPLEKTKSKAKGQPREPQVYTLPPSRDELTKNQVSLT 545

Db 302 TVLHODWLNGKEYKCKVSNKALPAPLEKTKSKAKGQPREPQVYTLPPSRDELTKNQVSLT 361  
QY 546 CLVKGFPYSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFSGS 605  
Db 362 CLVKGFPYSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFSGS 421  
QY 606 VMHEALHNHYTKSLSPG 625  
Db 422 VMHEALHNHYTKSLSPG 441

RESULT 87  
PCT-US96-10043-11  
Sequence 11, Application PC/TUS9610043  
GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation  
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES  
TITLE OF INVENTION: AND METHODS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/10043  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000,213  
FILING DATE: 14-JUN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lech, Karen F.  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 00786/284001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-10043-11

Query Match 35.6%; Score 1215.5; DB 5; Length 437;  
Best Local Similarity 97.0%; Pred. No. 2e-88;  
Matches 227; Conservative 0; Mismatches 2; Indels 5; Gaps 1;

QY 397 PAPEPKSCDKHTHC-----PELLGSPSVLFPKPDKDTLMISRTPEVTCVVDVSHEDPE 451  
Db 203 PEGEPKSCDKHTHCPCPAPAPLLGSPSVLFPKPDKDTLMISRTPEVTCVVDVSHEDPE 262  
QY 452 VRFNMYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHODWLNGKEYKCKVSNKALPAP 511  
Db 263 VRFNMYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHODWLNGKEYKCKVSNKALPAP 322  
QY 512 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYK 571  
Db 323 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYK 382

Qy 572 TTPVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTOKSLSPG 625  
Db 383 TTPVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTOKSLSPG 436

## RESULT 88

US-08-378-939-10  
Sequence 10, Application US/08378939  
Patent No. 5876961  
GENERAL INFORMATION:  
APPLICANT: CROME, JAMES SCOTT  
APPLICANT: LEWIS, ALAN PETER  
TITLE OF INVENTION: PRODUCTION OF ANTIBODIES  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROTHWELL, FIGG, ERNST & KIRZ  
STREET: 555 THIRTEENTH ST. N.W.  
CITY: WASHINGTON  
STATE: D. C.  
COUNTRY: U.S.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/378,939  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/952640  
FILING DATE: 01-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: ERNST, BARBARA G  
REGISTRATION NUMBER: 30,377  
REFERENCE/DOCKET NUMBER: 1808-118  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 783-6040  
TELEFAX: (202) 783-6031  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 476 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-378-939-10

Query Match 35.6%; Score 1215; DB 2; Length 476;

Best Local Similarity 97.8%; Pred. No. 2.5e-88;  
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

Qy 400 EPKSCDKTHTC-----PELLGSPVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKXF 454  
Db 245 EPKSCDKTHTCPCPAPELLGSPVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKXF 304  
Qy 455 NMVYDGVVNAKTPREEOYNSTRVVSUVTLVHODMLNGEKYCKSNKALPAPIEKT 514  
Db 305 NMVYDGVVNAKTPREEOYNSTRVVSUVTLVHODMLNGEKYCKSNKALPAPIEKT 364  
Qy 515 ISKAKGQREPOVYTLPPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTP 574  
Db 365 ISKAKGQREPOVYTLPPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTP 424  
Qy 575 PVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTOKSLSPG 625  
Db 425 PVLDSGSEFLYSKLTVDKSRMOQGNVSCSVMEHALHNHYTOKSLSPG 475

RESULT 89  
US-09-157-452B-12  
Sequence 12, Application US/09157452B  
Patent No. 6482409

GENERAL INFORMATION:  
APPLICANT: Lobb, Roy R.  
APPLICANT: Butkly, Linda C.  
TITLE OF INVENTION: TREATMENT FOR INFLAMMATORY BOWEL DISEASE  
FILE REFERENCE: 10274-004003  
CURRENT APPLICATION NUMBER: US/09/157,452B  
CURRENT FILING DATE: 1998-09-21  
PRIOR APPLICATION NUMBER: US 08/950,660  
PRIOR FILING DATE: 1997-10-15  
PRIOR APPLICATION NUMBER: US 08/373,857  
PRIOR FILING DATE: 1995-01-18  
PRIOR APPLICATION NUMBER: US 08/284,603  
PRIOR FILING DATE: 1994-08-11  
PRIOR APPLICATION NUMBER: PCT/US93/00924  
PRIOR FILING DATE: 1993-02-02  
PRIOR APPLICATION NUMBER: US 07/835,139  
PRIOR FILING DATE: 1992-02-12  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 12  
LENGTH: 446  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-157-452B-12

Query Match 35.5%; Score 1213.5; DB 4; Length 446;

Best Local Similarity 52.4%; Pred. No. 3e-88;  
Matches 270; Conservative 40; Mismatches 92; Indels 113; Gaps 14;

Qy 126 TANSDFHLQ-GSLTTLTLESPGSSP-----SVQCRSP-RGKNIOGG-KTLASVQLELQ 177  
Db 29 TTPSRVLAQGDGVSLSCTGTGCESPFPMRQIDBPLNGKYNBETSTLTNNPVSFG 88  
Qy 178 DSGTWCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVE-FSPPLAFTVEKLTGSG 236  
Db 89 NEHSYLTATCESRKLE-----KGIOVEIYSFP----- 116  
Qy 237 ELWMQARASSSKSWIFDKNKEVSVKRYTQDPKLOMGKKLPLHLTPQALPOYASGN 296  
Db 117 -----KDPRIHL-----SGPLAGRPITVKCSADVPP----- 144  
Qy 297 LTLALBAKTKLHOEVALVMRATOLQKNLTCEVWGPTSPMLSLKLENKAVSKREK 356  
Db 145 -----FRLRLIDLLKGDHLMKSG--EFLBDAKK-----SLFTKSLLEV----- 181  
Qy 357 PVMVLNPEAGMOCLLSDSQVLL-ESNIVLPTWSTPVCPCAPPEKSCDKTHTC----- 410  
Db 182 -----FTFVIDIGKVLVCRAKLHIDEMDSVPTVROAVKELQVDKTHCPCPCA 230  
Qy 411 PELLGSPVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKKNWYDGVVNAKTP 470  
Db 231 PELLGSPVFLPPPKKDTLMISRTPEVTCVVDVSHEDPEKKNWYDGVVNAKTP 290  
Qy 471 REEOYNSTRVVSUVTLVHODMLNGEKYCKSNKALPAPIEKTISKAKGQREPOVYTL 530  
Db 291 REEOYNSTRVVSUVTLVHODMLNGEKYCKSNKALPAPIEKTISKAKGQREPOVYTL 350  
Qy 531 PPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTPPVLDSGSEFLYSKLT 590  
Db 351 PPSRDELTKNOVSLTCLVGFYPSDIAVEMESNGQPENNYKTPPVLDSGSEFLYSKLT 410  
Qy 591 VDKSRMOQGNVSCSVMEHALHNHYTOKSLSPG 625  
Db 411 VDKSRMOQGNVSCSVMEHALHNHYTOKSLSPG 445

RESULT 90  
US-08-458-516-13  
Sequence 13, Application US/08458516  
Patent No. 5777085  
GENERAL INFORMATION:  
APPLICANT: Co, Man Sung  
APPLICANT: Tso, J. Yun

```

; TITLE OF INVENTION: Humanized Antibodies Reactive with
; TITLE OF INVENTION: GPIIb/IIIa
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,516
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/059,159
; FILING DATE: 03-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-37-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-458-516-13

Query Match      35.5%; Score 1213; DB 1; Length 449;
Best Local Similarity 97.8%; Pred. No. 3.3e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 400 EPPKSCDKHTTC-----PELGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 454
DB 218 EPPKSCDKHTTCPCPPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 277
QY 455 NMVYDGEVHNAAKTPREEQYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKT 514
DB 278 NMVYDGEVHNAAKTPREEQYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKT 337
QY 515 ISRAKQPREPOVYITLPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNQGPENNYKTP 574
DB 338 ISRAKQPREPOVYITLPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNQGPENNYKTP 397
QY 575 PVLDSGSFLLYSKLTVDKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 625
DB 398 PVLDSGSFLLYSKLTVDKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 448

RESULT 91
US-09-485-737B-67
; Sequence 67, Application US/09485737B
; Patent No. 6350860
; GENERAL INFORMATION:
; APPLICANT: Buysse, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC SHOCK.
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/09/485,737B
; PRIOR APPLICATION NUMBER: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
```

```

; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 67
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
; US-09-485-737B-67

Query Match      35.5%; Score 1212.5; DB 4; Length 468;
Best Local Similarity 45.3%; Pred. No. 3.9e-88;
Matches 287; Conservative 42; Mismatches 102; Indels 203; Gaps 17;

QY 11 LVLQALPAPATQGNKVLGKGGDTVELCTASQKSIQFHKNSNQIKLQNGSFLT 70
DB 17 VLISQVQVQSGSE-----LKKPGASVKISCKAS--GTFITDYGKMWYKQAPGQ---L 65
QY 71 KQPSKLNDRADSRRLMD-QGNFP-----LIINKLIEDSDYICEVEDQKEV 118
DB 66 KMWGMINTYGTSTYVDPKGRFVSLDTSVSAAYLQISLKAEDTATVFC----- 116
QY 119 QLLVGLTANSQTHLQ--GQSLTLTLESPRSSSPVQGRSFRGKRIQGGKTLTSVQL 176
DB 117 -----ARRGFYADYWGQGTITVVSASATKGPVFPPLAPSSKSTSG--TVALGCL-- 165
QY 177 QDSGTWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPLAFVTEKLTGSG 236
DB 166 -----VKKYFPEPVYS----- 177
QY 237 ELWQAEKASSSSKSWITFDLKNKEVSVKRYTQPKLQMGKULPLHLTPQALPQVAGSN 296
DB 178 ---WNSGALTSG-----VH--TPPAVL-QSSGLYS 201
QY 297 LTLAEAKTGKLGHEVNLVVMRATQKRLTCEWQPTSPKMLSLKENKAKSKRER 356
DB 202 LSSVYVTPSSSLGTQYI-----CNV-----NHRKSNTKVKD 233
QY 357 PWWNLNPEAGMMQCLLSDSGQVLLBSNIKVLPWTSTPVPAPPEPKSCDKHTTC-----P 411
DB 234 RV-----EPKSCDKHTTCPCPPAP 252
QY 412 ELGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTPR 471
DB 253 ELGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTPR 312
QY 472 EEQYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKTISAKQPREPOVYITP 531
DB 313 EEQYNSTYRVSVLYLVHODMNGKEYCKVSNKALPAPIEKTISAKQPREPOVYITP 372
QY 532 PSRDELTKNOVSLTCLVKGFYPSDIAVEMESNQGPENNYKTPPVLDSDGSFLLYSKLT 591
DB 373 PSREMTKNOVSLTCLVKGFYPSDIAVEMESNQGPENNYKTPPVLDSDGSFLLYSKLT 432
QY 592 DKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 625
DB 433 DKSRMOQGNVFCSVHMEALHNHYTQKSLSLSPG 466

RESULT 92
US-09-485-737B-90
; Sequence 90, Application US/09485737B
; Patent No. 6350860
; GENERAL INFORMATION:
; APPLICANT: Buysse, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC SHOCK.
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
```

```

; CURRENT APPLICATION NUMBER: US/09/485.737B
; CURRENT FILING DATE: 2000-02-14
; PCT REFERENCE: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 96870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
US-09-485-737B-90
```

```

Query Match          35.5%; Score 1212.5; DB 4; Length 711;
Best Local Similarity 45.3%; Pred. No. 7.1e-88;
Matches 287; Conservative 42; Mismatches 102; Indels 203; Gaps 17;
```

```

QY 11 LVLVLALLPRAATGKNNVLAGKDDVLELTCTASQKSIQFMKNSQIKILNQSGFLLT 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 17 VILSQVQLVQSGSE-----LKKPGASVXISCKAS---GYFTDYGMWVVKQAPGQ---L 65

QY 71 KGPBKLDRAADRSRLMD--QGNFP-----LIIKNLKITDSDTYICEVEDQKEV 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 66 KMMGWINTYTGESTYVDKGRFVFSLDTSVAAYLIQISLKADDTATYFC----- 116

QY 119 QLVFGLTANSDTHLQ--GQSLTLTLESPPGSSPVOCRRPGKNIQGGKTLISVSOLE 176
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 117 -----ARGFAMDYWGCGTIVTVSSASTKPSVFPPLAPSSKISG--TALAGCL-- 165

QY 177 QDSGTWTCTYQNKQKVEFKIDIVLAFQKASIVYKKEGQVFSFPLAFTVEKLTGSG 236
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 166 -----VADYFPEPVTVS----- 177

QY 237 ELWWQABRASSKSWITFDLKNKEVSQKVTQDRLQMGKLPILHLPLPOLPYAGSGN 296
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 178 ---WNSGALTSG-----VH--TFPAVL--QSSGLYS 201

QY 297 LTLALEAKTGKLEHVENLVVMRATOLQNLTCFVWGPFSPLMLSLKLENKAVSKREK 356
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 202 LSSVVTYPPSSSLGTQTYI-----CNV-----NHKPSNTKVDK 233

QY 357 PVWVLNBEAGMQLLSDSGVLLSENIKVLPTWSTPVCPAPRPSKCDKTHTC-----P 411
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 234 RV-----EPKSCDKTHTCPCPAP 252

QY 412 ELLEGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNVYDGEVHNAAKTFR 471
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 253 ELLEGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNVYDGEVHNAAKTFR 312

QY 472 EEOYNSTYRVVSVTLVHODMNLNGEKYCKVSNKALPAPIEKTIKAQGRREOVYTLR 531
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 313 EEOYNSTYRVVSVTLVHODMNLNGEKYCKVSNKALPAPIEKTIKAQGRREOVYTLR 372

QY 532 PSRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPPVLDSDGSFFLYSKLTV 591
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 373 PSRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPPVLDSDGSFFLYSKLTV 432

QY 592 DKSRWQGNVFGSCVMHIALHNHYTQKSLSLSPG 625
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 433 DKSRWQGNVFGSCVMHIALHNHYTQKSLSLSPG 466
```

```

RESULT 93
US-09-313-942-7
; Sequence 7, Application US/09313942
; Patent No. 6472179
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
```

```

; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING
; FILE REFERENCE: REG 203-A
; CURRENT APPLICATION NUMBER: US/09/313.942
; CURRENT FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 09/313,942
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/101,858
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-313-942-7
```

```

Query Match          35.5%; Score 1212.5; DB 4; Length 859;
Best Local Similarity 93.1%; Pred. No. 9.3e-88;
Matches 229; Conservative 3; Mismatches 5; Indels 9; Gaps 2;
```

```

QY 389 TWSTPV---PCPAPRPSKCDKTHTC-----PELIGPSVFLFPKPKDTLMISRTPEVT 439
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 607 TFFTTPKPAQGEIISGSEPKSCDKTHTCPCPAPRPSVFLFPKPKDTLMISRTPEVT 666

QY 440 CVVVDVSHEDPEVKFNNVYDGEVHNAAKTFRREOVNSTYRVVSVTLVHODMNLNGEKYK 499
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 667 CVVVDVSHEDPEVKFNNVYDGEVHNAAKTFRREOVNSTYRVVSVTLVHODMNLNGEKYK 726

QY 500 CKVSNKALPAPIEKTIKAQGRREOVYTLPSRDELTKNOVSLTCLVKGFPYPSDIAVE 559
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 727 CKVSNKALPAPIEKTIKAQGRREOVYTLPSRDELTKNOVSLTCLVKGFPYPSDIAVE 786

QY 560 MESNGQENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFGSCVMHIALHNHYTQKS 619
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 787 MESNGQENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFGSCVMHIALHNHYTQKS 846

QY 620 LSLSPG 625
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 847 LSLSPG 852
```

```

RESULT 94
US-09-313-942-9
; Sequence 9, Application US/09313942
; Patent No. 6472179
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING
; FILE REFERENCE: REG 203-A
; CURRENT APPLICATION NUMBER: US/09/313.942
; CURRENT FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 09/313,942
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: 60/101,858
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-313-942-9
```

```

Query Match          35.5%; Score 1212.5; DB 4; Length 951;
Best Local Similarity 97.8%; Pred. No. 1.1e-87;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

```

QY 400 EPKSCDKTHTC-----PELIGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 454
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 720 EPKSCDKTHTCPCPAPRPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 779
```

QY 455 NMVVDGEVHNNAKTPREEOYNSTYRVSVLTVLHODMLNGKEYCKCKVSKALPAPIEKT 514  
| | | | |  
DB 780 NMVVDGEVHNNAKTPREEOYNSTYRVSVLTVLHODMLNGKEYCKCKVSKALPAPIEKT 839  
| | | | |  
QY 515 ISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFSYPSDIAVEMESNQEPENNYKTP 574  
| | | | |  
DB 840 ISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFSYPSDIAVEMESNQEPENNYKTP 899  
| | | | |  
QY 575 PVLDSGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALHNYTQKSLSLSPG 625  
| | | | |  
DB 900 PVLDSGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALHNYTQKSLSLSPG 950  
| | | | |

## RESULT 95

US-09-247-352-3  
; Sequence 3, Application US/09247352  
; Patent No. 6312693  
; GENERAL INFORMATION:  
; APPLICANT: Aruffo, Alejandro A.  
; APPLICANT: Siadak, Anthony W.  
; APPLICANT: Berry, Karen K.  
; APPLICANT: Harris, Linda  
; APPLICANT: Thorne, Barbara A.  
; APPLICANT: Bajorath, Jurgen  
; APPLICANT: Huse, William D.  
; APPLICANT: Wu, Herren  
; APPLICANT: Watkins, Jeffrey D.  
; TITLE OF INVENTION: ANTIBODIES AGAINST HUMAN CD40  
; FILE REFERENCE: DB2A SEQUENCE  
; CURRENT APPLICATION NUMBER: US/09/247,352  
; EARLIER APPLICATION NUMBER: 1999-02-10  
; EARLIER FILING DATE: 1998-02-19  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 451  
; TYPE: PRT  
; ORGANISM: Human and Mouse  
US-09-247-352-3

Query Match 35.5%; Score 1212; DB 4; Length 451;  
Best Local Similarity 46.7%; Pred. No. 4e-88;  
Matches 286; Conservative 34; Mismatches 103; Indels 190; Gaps 17;  
QY 30 LKKGGDTVELTCTAQSQKSIQFHWNKSNQIKILGNQGSFLTKGPSKLNDRADSRSLMD- 88  
| | | | |  
DB 11 LKKGGDTVELTCTAQSQKSIQFHWNKSNQIKILGNQGSFLTKGPSKLNDRADSRSLMD- 88  
| | | | |  
QY 89 QGNFP-----LITIKLIKEDSDTYICEVEDQKEVOULVFGLTANSPTHLLOGQ 137  
| | | | |  
DB 65 KGRFASLETSANTAYLQISNLKNEBTATYFC-VRSNGNAYDLAYFA-----YWGQ 114  
| | | | |  
QY 138 SLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTWCTVLOKQKVEFKI 197  
| | | | |  
DB 115 GLTVTVAASATKGSVFPLAPSSKSTSG--TALAGCL----- 150  
| | | | |  
QY 198 DIVVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMWQAERASSSKSWITFDLK 257  
| | | | |  
DB 151 -----VKDYFPEPVTVS-----WNSGALTSG----- 171  
| | | | |  
QY 258 NKEVSVRVYTOPDKLQMGKKLPLHLTLPOALPOYAGSGLTLALEATGKLHQBENVLVVM 317  
| | | | |  
DB 172 -----VH-TFPAVL-QSSGLYSLSSVTVTPSSSLGTQTYI--- 204  
| | | | |  
QY 318 RATOLQKNLTCCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLANPEAGMWOCCLSDSGQ 377  
| | | | |  
DB 205 -----CNV-----NHKPSNTKYDKV----- 220  
| | | | |  
QY 378 VLESNIKVLPTWSTVPCPAPBPKSCDKTHTC-----PELLGSPSVFLFPKPKDTLMT 432  
| | | | |  
DB 221 -----EPKSCDKTHTCPPCAPPELLGSPSVFLFPKPKDTL-1 257  
| | | | |

QY 433 SRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEOYNSTYRVSVLTVLHODM 492  
| | | | |  
DB 258 SRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEOYNSTYRVSVLTVLHODM 317  
| | | | |  
QY 493 LNKKEYCKVSKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFSY 552  
| | | | |  
DB 318 LNKKEYCKVSKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFSY 377  
| | | | |  
QY 553 PBDIAVEMESNQEPENNYKTPPVLDSGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALH 612  
| | | | |  
DB 378 PBDIAVEMESNQEPENNYKTPPVLDSGSPFLYSKLTVDKSRMOQGNVFCSCVMHEALH 437  
| | | | |  
QY 613 NHYTQKSLSLSPG 625  
| | | | |  
DB 438 NHYTQKSLSLSPG 450  
| | | | |

## RESULT 96

US-09-466-635-3  
; Sequence 3, Application US/09466635  
; Patent No. 6413514  
; GENERAL INFORMATION:  
; APPLICANT: Aruffo, Alejandro A.  
; APPLICANT: Siadak, Anthony W.  
; APPLICANT: Berry, Karen K.  
; APPLICANT: Harris, Linda  
; APPLICANT: Thorne, Barbara A.  
; APPLICANT: Bajorath, Jurgen  
; TITLE OF INVENTION: ANTIBODIES AGAINST HUMAN CD40  
; FILE REFERENCE: DB2 SEQUENCE  
; CURRENT APPLICATION NUMBER: US/09/466,635  
; CURRENT FILING DATE: 1999-12-17  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 451  
; TYPE: PRT  
; ORGANISM: Human and Mouse  
US-09-466-635-3

Query Match 35.5%; Score 1212; DB 4; Length 451;  
Best Local Similarity 46.7%; Pred. No. 4e-88;  
Matches 286; Conservative 34; Mismatches 103; Indels 190; Gaps 17;  
QY 30 LKKGGDTVELTCTAQSQKSIQFHWNKSNQIKILGNQGSFLTKGPSKLNDRADSRSLMD- 88  
| | | | |  
DB 11 LKKGGDTVELTCTAQSQKSIQFHWNKSNQIKILGNQGSFLTKGPSKLNDRADSRSLMD- 88  
| | | | |  
QY 89 QGNFP-----LITIKLIKEDSDTYICEVEDQKEVOULVFGLTANSPTHLLOGQ 137  
| | | | |  
DB 65 KGRFASLETSANTAYLQISNLKNEBTATYFC-VRSNGNAYDLAYFA-----YWGQ 114  
| | | | |  
QY 138 SLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDSDGTWCTVLOKQKVEFKI 197  
| | | | |  
DB 115 GLTVTVAASATKGSVFPLAPSSKSTSG--TALAGCL----- 150  
| | | | |  
QY 198 DIVVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMWQAERASSSKSWITFDLK 257  
| | | | |  
DB 151 -----VKDYFPEPVTVS-----WNSGALTSG----- 171  
| | | | |  
QY 258 NKEVSVRVYTOPDKLQMGKKLPLHLTLPOALPOYAGSGLTLALEATGKLHQBENVLVVM 317  
| | | | |  
DB 172 -----VH-TFPAVL-QSSGLYSLSSVTVTPSSSLGTQTYI--- 204  
| | | | |  
QY 318 RATOLQKNLTCCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLANPEAGMWOCCLSDSGQ 377  
| | | | |  
DB 205 -----CNV-----NHKPSNTKYDKV----- 220  
| | | | |  
QY 378 VLESNIKVLPTWSTVPCPAPBPKSCDKTHTC-----PELLGSPSVFLFPKPKDTLMT 432  
| | | | |  
DB 221 -----EPKSCDKTHTCPPCAPPELLGSPSVFLFPKPKDTL-1 257  
| | | | |  
QY 433 SRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREEOYNSTYRVSVLTVLHODM 492  
| | | | |

Db 258 SRTEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKRREQYSTYRVSVLTFLHQM 317  
Qy 493 LINGEYKCKVKNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVNGFY 552  
Db 318 LINGEYKCKVKNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVNGFY 377  
Qy 553 PSDIAYEVESNGQPENNYKTPPVLDSDGSFFLYSKLTVDSKRMQGNVFCSCVMHEALH 612  
Db 378 PSDIAYEVESNGQPENNYKTPPVLDSDGSFFLYSKLTVDSKRMQGNVFCSCVMHEALH 437  
Qy 613 NHYTQKSLISLSPG 625  
Db 438 NHYTQKSLISLSPG 450  
RESULT 97  
US-09-301-593-30  
; Sequence 30, Application US/09301593A  
; Patent No. 6455677  
; GENERAL INFORMATION:  
; APPLICANT: Park, John E.  
; APPLICANT: Garin-Chesa, Pilar  
; APPLICANT: Bamberger, Uwe  
; APPLICANT: Legier, Olivier  
; APPLICANT: Saldanha, Jose W.  
; APPLICANT: Rettig, Wolfgang J.  
; TITLE OF INVENTION: FAP-specific Antibody with Improved Producibility  
; FILE REFERENCE: 0652.1890001  
; CURRENT APPLICATION NUMBER: US/09/301.593A  
; CURRENT FILING DATE: 1999-04-29  
; EARLIER APPLICATION NUMBER: EP 98107925.4  
; EARLIER FILING DATE: 1998-04-30  
; EARLIER APPLICATION NUMBER: US 60/086.049  
; EARLIER FILING DATE: 1998-05-18  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 30  
; LENGTH: 472  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-301-593-30  
Query Match 35.5%; Score 1212; DB 4; Length 472;  
Best Local Similarity 49.4%; Pred. No. 4.3e-88;  
Matches 266; Conservative 39; Mismatches 74; Indels 160; Gaps 13;  
Qy 147 PGSSPSVQCRRSPR-----GKNIOG-----G 166  
Db 33 PGASVKNKCKTSRTFTFTYTHWVRQSHGKSLMWIGINPNNGIPIYNNQKFKGRATLTVG 92  
Qy 167 KTLVSQLEL-----QDSGTWTCVLYNQKVEFKIDIVLAFOKASSIYKKEGEOYEF 221  
Db 93 KSSSTAYVELRSLTSEDSAVYFCA---RRRIAYGYD-----EGHAMDY 132  
Qy 222 SFPLAFTYKLTGSGELMWQAEARASSSKSWITPDKNKEVSVKRVTQDPKLOMGKULPLH 281  
Db 133 -----WGQSTVTVSSS-----TKGPSVFPLAPSSKTSQGTALG 168  
Qy 282 LTLPLALPOYA-----GSGNLTLLALEA-----KTGKLHQBVLVVMRATQL-QKNLTCEVW 331  
Db 169 CLVNDYFPEPVTAVSNKSLTSGVHTPFAVLQSSGLYSLSSVTVYPSLSLGTQYIYICV- 227  
Qy 332 GPSPKMLSLKLENKEAKVSKREKPVAVLPEAGMOCCLISDGOVLLESNIVLPTWS 391  
Db 228 -----NHKPSNTKVKDKV----- 240  
Qy 392 TPVCPAPPEKSCOKTHTC-----PELLGSPVFLPPKPKDTLMISTPEVTCVVVDVS 446  
Db 241 -----EPKSCDTHTCPPCPAPPELLGSPSVFLPPKPKDTLMISTPEVTCVVVDVS 292  
Qy 447 HEDPEVKFNWYVDGVEVHNAKTKRREQYSTYRVSVLTFLHQM LINGEYKCKVKNK 506

Db 293 HEDPEVKFNWYVDGVEVHNAKTKRREQYSTYRVSVLTFLHQM LINGEYKCKVKNK 352  
Qy 507 LPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVNGFYPSDIAYEVESNGQ 566  
Db 353 LPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVNGFYPSDIAYEVESNGQ 412  
Qy 567 ENNYKTPPVLDSDGSFFLYSKLTVDSKRMQGNVFCSCVMHEALHNYHTQKSLISLSPG 625  
Db 413 ENNYKTPPVLDSDGSFFLYSKLTVDSKRMQGNVFCSCVMHEALHNYHTQKSLISLSPG 471  
RESULT 98  
US-08-595-043A-50  
; Sequence 50, Application US/08595043A  
; Patent No. 5935824  
; GENERAL INFORMATION:  
; APPLICANT: SGARLATO, GREGORY D.  
; TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/595.043A  
; FILING DATE: 31-JAN-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: SGAR-00371  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-595-043A-50  
Query Match 35.5%; Score 1211.5; DB 2; Length 232;  
Best Local Similarity 97.8%; Pred. No. 1.7e-88;  
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;  
Qy 400 EPKSCDTHTC-----PELLGSPVFLPPKPKDTLMISTPEVTCVVVDVSHEDPEVK 454  
Db 1 EPKSCDTHTCPPCPAPPELLGSPSVFLPPKPKDTLMISTPEVTCVVVDVSHEDPEVK 60  
Qy 455 NMVYDGEVHNAKTKRREQYSTYRVSVLTFLHQM LINGEYKCKVKNKALPAPIETK 514  
Db 61 NMVYDGEVHNAKTKRREQYSTYRVSVLTFLHQM LINGEYKCKVKNKALPAPIETK 120  
Qy 515 ISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVNGFYPSDIAYEVESNGQPENNYKTP 574  
Db 121 ISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVNGFYPSDIAYEVESNGQPENNYKTP 180  
Qy 575 PVLDSDGSFFLYSKLTVDSKRMQGNVFCSCVMHEALHNYHTQKSLISLSPG 625  
Db 181 PVLDSDGSFFLYSKLTVDSKRMQGNVFCSCVMHEALHNYHTQKSLISLSPG 231  
RESULT 99  
US-09-178-869-2

```
; Sequence 2, Application US/09178869B
; Patent No. 6197294
; GENERAL INFORMATION:
; APPLICANT: Tao, Weng
; APPLICANT: Wong, Shou
; APPLICANT: Hickey, William F.
; APPLICANT: Hammang, Joseph P.
; APPLICANT: Baetge, E. Edward
; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION
; FILE REFERENCE: 17810-043
; CURRENT APPLICATION NUMBER: US/09/178,869B
; CURRENT FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-178-869-2
```

```
Query Match          35.5%; Score 1211.5; DB 3; Length 331;
Best Local Similarity 97.8%; Pred. No. 2.8e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

```
QY 400 EPKSCDKTHTC-----PELLGSPSVFLFPPKPKDTLMIISRTPEVTCVVDVSHEDPEVKF 454
DB 100 EPKSCDKTHTCPCPAPABELLGSPSVFLFPPKPKDTLMIISRTPEVTCVVDVSHEDPEVKF 159
QY 455 NMVYDGEVHNAAKTREBOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 514
DB 160 NMVYDGEVHNAAKTREBOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 219
QY 515 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 574
DB 220 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 279
QY 575 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
DB 280 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 330
```

```
RESULT 100
US-09-761-413-2
; Sequence 2, Application US/09761413
; Patent No. 6506891
; GENERAL INFORMATION:
; APPLICANT: Tao, Weng
; APPLICANT: Wong, Shou
; APPLICANT: Hickey, William F.
; APPLICANT: Hammang, Joseph P.
; APPLICANT: Baetge, E. Edward
; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION
; FILE REFERENCE: 17810-043
; CURRENT APPLICATION NUMBER: US/09/761,413
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US/09/178,869
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-761-413-2
```

```
Query Match          35.5%; Score 1211.5; DB 4; Length 331;
Best Local Similarity 97.8%; Pred. No. 2.8e-88;
Matches 226; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

```
QY 400 EPKSCDKTHTC-----PELLGSPSVFLFPPKPKDTLMIISRTPEVTCVVDVSHEDPEVKF 454
DB 100 EPKSCDKTHTCPCPAPABELLGSPSVFLFPPKPKDTLMIISRTPEVTCVVDVSHEDPEVKF 159
```

```
QY 455 NMVYDGEVHNAAKTREBOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 514
DB 160 NMVYDGEVHNAAKTREBOYNSTYRVSVLTVLHODMLNGKEYCKCKVSNKALPAPIEKT 219
QY 515 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 574
DB 220 ISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTT 279
QY 575 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
DB 280 PVLDSGSEFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 330
```

```
Search completed: August 3, 2004, 13:17:11
Job time : 29.0695 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:50:09 ; Search time 65.926 Seconds  
(without alignments)  
2777.216 Million cell updates/sec

Title: SEQ7  
Perfect score: 3414  
Sequence: 1 MNRGVPRRLHLVLQLALP.....DETCAEQDELGLWTTDP 648

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 125 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3239.5	94.6	631	1 AAP93009	AAP93009 Genetic C
2	3239.5	94.6	631	3 AAB19508	AAB19508 CD4-IgG1
3	3229.5	94.6	631	3 AAY51079	AAY51079 Human fus
4	3229.5	94.6	631	3 AAY59169	AAY59169 CD4-Ig fu
5	3176.5	93.0	729	1 AAP93008	AAP93008 Genetic C
6	3176.5	93.0	729	3 AAB19507	AAB19507 CD4-IgG1
7	3176.5	93.0	729	3 AAY59168	AAY59168 CD4-Ig fu
8	3166.5	92.8	729	3 AAY51078	AAY51078 Human fus
9	3163	92.6	616	3 AAY51082	AAY51082 Human fus
10	3163	92.6	616	3 AAY59172	AAY59172 CD4-Ig fu
11	3132	91.7	614	1 AAP93012	AAP93012 Genetic C
12	3071	90.0	616	3 AAB19511	AAB19511 CD4-IgG1
13	2212	64.8	534	2 AAR26531	AAR26531 Sequence
14	2140	62.7	435	2 AAR26530	AAR26530 Sequence
15	2116	62.0	530	2 AAR26783	AAR26783 CD4-IgG2
16	2116	62.0	530	3 AAY85080	AAY85080 CD4-IgG2
17	2116	62.0	530	4 AAB67323	AAB67323 CD4-IgG2
18	2116	62.0	530	4 AAB80884	AAB80884 Human CD4
19	2116	62.0	530	6 ABG71123	ABG71123 CD4-immun
20	2107.5	61.7	450	6 AAE37574	AAE37574 Human DID
21	2106	61.7	530	6 AAR46679	AAR46679 CD4-IgG2
22	2088	61.2	449	6 AAE37576	AAE37576 Human FDI
23	2073	60.7	481	3 AAB19510	AAB19510 CD4-IgG1 f
24	2073	60.7	481	3 AAY59171	AAY59171 CD4-Ig fu
25	2067	60.5	481	3 AAY51081	AAY51081 Human fus

26	2065	60.5	481	1 AAP93011	AAP93011 Genetic C
27	2059	60.3	432	2 AAR26782	AAR26782 CD4-gamma
28	2059	60.3	432	2 AAR46678	AAR46678 CD4-gamma
29	2059	60.3	432	2 AAY85079	AAY85079 Human CD4
30	2059	60.3	432	4 AAB67322	AAB67322 CD4-gamma
31	2059	60.3	432	4 AAB80883	AAB80883 Human CD4
32	2059	60.3	432	6 ABG71122	ABG71122 CD4-gamma
33	2046.5	59.9	436	3 AAY51080	AAY51080 Human fus
34	2041	59.8	474	3 AAY59170	AAY59170 CD4-Ig fu
35	2036.5	59.7	532	2 AAR27278	AAR27278 CD4-gamma
36	2036.5	59.7	532	2 AAR78678	AAR78678 T-cell re
37	2036.5	59.7	532	2 AAR89458	AAR89458 CD4-eta f
38	2036.5	59.7	532	2 AAR89458	AAR89458 CD4-eta f
39	2036.5	59.7	575	2 AAR27276	AAR27276 CD4-zeta
40	2036.5	59.7	575	2 AAR8676	AAR8676 T-cell re
41	2036.5	59.7	575	2 AAR89456	AAR89456 CD4-zeta
42	2036.5	59.7	575	2 AAW02213	AAW02213 CD4-T-cel
43	2036.5	59.5	575	2 AAW83140	AAW83140 Chimeric
44	2032.5	59.5	462	2 AAR27277	AAR27277 CD4-eta p
45	2032.5	59.5	462	2 AAR78677	AAR78677 T-cell re
46	2032.5	59.5	462	2 AAR89457	AAR89457 CD4-gamma
47	2032.5	59.5	462	2 AAW02214	AAW02214 CD4-Fc re
48	2032.5	59.5	462	2 AAW83142	AAW83142 Chimeric
49	2032.5	59.5	532	2 AAW02215	AAW02215 CD4-T-cel
50	2029	59.4	398	2 AAR89450	AAR89450 CD4-D1-D4
51	2029	59.4	458	2 AAR07769	AAR07769 DNA encod
52	2026	59.3	398	2 AAR78673	AAR78673 CD4 domai
53	2026	59.3	416	3 AAB19509	AAB19509 CD4-IgM f
54	2021	59.2	400	2 AAR06374	AAR06374 Truncated
55	2021	59.2	458	1 AAB81990	AAB81990 Clone PT4
56	2021	59.2	458	1 AAP91369	AAP91369 T4 protei
57	2021	59.2	458	2 AAY39826	AAY39826 Soluble h
58	2021	59.2	458	2 AAR04032	AAR04032 Full leng
59	2021	59.2	2050	2 AAR07641	AAR07641 Deduced s
60	2018	59.1	394	1 AAP93506	AAP93506 Derived s
61	2017	59.1	402	1 AAP91922	AAP91922 Sequence
62	2017	59.1	402	1 AAP94757	AAP94757 Sequence
63	2015	59.0	394	3 AAY88328	AAY88328 T4 glycop
64	2015	59.0	458	3 AAY88329	AAY88329 T4 glycop
65	2015	59.0	458	4 AAB81502	AAB81502 Human CD4
66	2015	59.0	458	7 ADD25609	ADD25609 Binding d
67	2015	59.0	458	7 ADE57489	ADE57489 Human Pro
68	2015	59.0	473	7 ADA44807	ADA44807 CD4/TCR C
69	2015	59.0	519	2 AAR20152	AAR20152 Human CD4
70	2013	59.0	458	2 AAR13491	AAR13491 Human CD4
71	2012	58.9	394	2 AAY39825	AAY39825 Soluble h
72	2007	58.8	458	4 AAG79087	AAG79087 Amino aci
73	2001	58.6	394	3 AAB07768	AAB07768 The solub
74	2001	58.5	458	7 ADE58841	ADE58841 Human CD4
75	1993	58.4	458	2 AAR06373	AAR06373 T4 encode
76	1993	58.4	524	1 AAP94703	AAP94703 Sequence
77	1993	58.4	2458	2 AAR07640	AAR07640 Deduced p
78	1993	58.3	2458	1 AAR04031	AAR04031 Full leng
79	1991	58.3	399	2 AAR20151	AAR20151 Chimpanze
80	1991	58.3	458	2 AAR11285	AAR11285 gp120 bin
81	1991	58.3	458	2 AAR10988	AAR10988 Chimpanze
82	1986	58.2	400	2 AAR20150	AAR20150 Chimpanze
83	1985	58.1	458	2 AAR04910	AAR04910 T4 protei
84	1982.5	57.0	399	1 AAP93010	AAP93010 Genetic C
85	1946	57.0	400	2 AAR20154	AAR20154 Sol. rhes
86	1929.5	56.5	729	2 AAR41042	AAR41042 CD4-GBPH
87	1921	56.3	458	2 AAR15149	AAR15149 CD4 coord
88	1908.5	55.9	942	2 AAR11041	AAR11041 CD4-GBP13
89	1908	55.9	384	1 AAP90833	AAP90833 Amino aci
90	1908	55.8	1786	2 AAR41043	AAR41043 CD4-EBB17
91	1904	55.8	434	1 AAP96151	AAP96151 Sequence
92	1904	55.8	434	1 AAP93557	AAP93557 Fusion of
93	1894	55.5	375	2 AAR07721	AAR07721 Recombina
94	1892	55.4	433	2 AAW41376	AAW41376 Human CD4
95	1891	55.4	369	2 AAY39824	AAY39824 Soluble h
96	1891	55.4	369	2 AAY88327	AAY88327 T4 glycop
97	1891	55.4	370	1 AAP93528	AAP93528 Human sol
98	1891	55.4	370	4 AAB83356	AAB83356 Human CD4

99 1887 55.3 433 3 AAY54500 Aay54500 Amino ac  
 100 1884 55.2 432 2 AAR74222 Aar74222 Epitope o  
 101 1884 55.2 432 2 AAY30514 Aay30514 Predicted  
 102 1884 55.2 432 6 ADA25188 Ada25188 CD4 epit  
 103 1879 55.0 370 2 AAR12956 Aar12956 Lys (64) C  
 104 1878 55.0 370 2 AAR12954 Aar12954 His (40) C  
 105 1878 55.0 370 2 AAR12964 Aar12964 Val (143)  
 106 1878 55.0 370 2 AAR12965 Aar12965 Arg (150)  
 107 1877 55.0 370 2 AAR12958 Aar12958 Gln (80) C  
 108 1874 54.9 370 2 AAR12951 Aar12951 Glu (17)  
 109 1872 54.8 370 2 AAR12957 Aar12957 Asn (72)  
 110 1871 54.8 370 2 AAR12959 Aar12959 Asn (88)  
 111 1871 54.7 370 2 AAR12962 Aar12962 Leu (127)  
 112 1866 54.7 370 2 AAR12966 Aar12966 Thr (162)  
 113 1865 54.6 370 2 AAR12952 Aar12952 Ile (23)  
 114 1864 54.6 370 2 AAR12967 Aar12967 Asp (155)  
 115 1862 54.5 370 2 AAR12955 Aar12955 Gly (48)  
 116 1861 54.5 370 2 AAR12963 Aar12963 His (132)  
 117 1861 54.5 370 2 AAR12960 Aar12960 Lys (99)  
 118 1861 54.5 435 1 AAR90992 Aar90992 Human CD4  
 119 1858 54.4 435 1 AAR12961 Aar12961 Ser (121)  
 120 1854 54.3 435 1 AAR12961 Aar12961 Ser (121)  
 121 1852 54.2 370 2 AAR12959 Aar12959 Sequence  
 122 1850 54.2 433 2 AAR08335 Aar08335 Thr (27)  
 123 1849 54.2 400 2 AAR08335 Aar08335 CD4 . 3/20  
 124 1848 54.1 458 2 AAR10987 Aar10987 Sol. rhes  
 125 1847 54.1 458 2 AAR20148 Aar20148 Rhesus mo  
 Aar20148 Cynomolgus

## ALIGNMENTS

RESULT 1  
 ID AAR93009 standard; protein; 631 AA.  
 XX AAR93009;  
 AC AAR93009;  
 DT 25-MAR-2003 (revised)  
 DT 02-NOV-1992 (first entry)  
 XX  
 DE Genetic construct which encodes CD4 linked to human IgG1 at the Esp site  
 DE upstream of the hinge region (fusion protein CD4E-gamma-1).  
 DE  
 XX Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy;  
 KM diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPJ25262-A.  
 XX  
 PD 26-JUL-1989.  
 XX  
 PF 20-JAN-1989; 89EP-00100913.  
 XX  
 PR 22-JAN-1988; 88US-00147351.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Seed B;  
 XX  
 DR WPI; 1989-214472/30.  
 DR N-PSDB; AAN90357.  
 XX  
 PT Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV  
 PT infections or detecting HIV or SIV in sample.  
 XX  
 PS Example; Table 2, Page 24-33; 68pp; English.  
 XX  
 CC The fusion protein genes of the invention pref. comprise cDNA sequences  
 CC which encode CD4 or a fragment which binds gp120 ligated to an expression  
 CC plasmid which encodes an antibody in which the variable region of the  
 CC gene has been deleted (see WO87-02671). The CD4 portion of the fusion

CC protein may comprise the complete CD4 sequence, the 370 AA extracellular  
 CC region and the membrane spanning domain, or the extracellular region. The  
 CC Ig heavy chain is pref. from IgM, IgG1 or IgG3. The following are  
 CC specifically claimed: fusion proteins CD4H-gamma-1, CD4Hmu, CD4Epmu, CD4E-  
 CC gamma1, and CD4Hmu (No. 67608), pCD4P-gamma1, pCD4E-gamma-1  
 CC (No. 67610). The plasmid containing (pCD4E-gamma-1) has been deposited in  
 CC E. coli (MC1061/p3) at the ATCC under accession number 67610. (Updated on  
 CC 25-MAR-2003 to correct PA field.)

XX Sequence 631 AA;

Query Match 94.6%; Score 3229.5; DB 1; Length 631;  
 Best Local Similarity 98.3%; Pred. No. 36-166;  
 Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;

QY 1 MNRGVFRLHLVLTALPAATGKVVYLGKKGTVELTCTASQKSIQFMKNSQIK 60  
 DB 1 MNRGVFRLHLVLTALPAATGKVVYLGKKGTVELTCTASQKSIQFMKNSQIK 60  
 QY 61 IIGNGSFLTKGSPKLNDRSRRLMDGNEPLIIKNKIDSDTYICEVEDQKEEVL 120  
 DB 61 IIGNGSFLTKGSPKLNDRSRRLMDGNEPLIIKNKIDSDTYICEVEDQKEEVL 120  
 QY 121 LVFGLTANSDDLQGSLLTLTLESPPGSSPEVQCRSPRGKNIQGGKTLVSQLEIDSG 180  
 DB 121 LVFGLTANSDDLQGSLLTLTLESPPGSSPEVQCRSPRGKNIQGGKTLVSQLEIDSG 180  
 QY 121 LVFGLTANSDDLQGSLLTLTLESPPGSSPEVQCRSPRGKNIQGGKTLVSQLEIDSG 180  
 DB 121 LVFGLTANSDDLQGSLLTLTLESPPGSSPEVQCRSPRGKNIQGGKTLVSQLEIDSG 180  
 QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKGEOVESFPPLAFYTEKLTGSGELMW 240  
 DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKGEOVESFPPLAFYTEKLTGSGELMW 240  
 QY 241 QAEPRSSSKSWITFDLKNKEVSKVKTODPKLOMKKPLMHLTLPOLQVYSGSLTLTA 300  
 DB 241 QAEPRSSSKSWITFDLKNKEVSKVKTODPKLOMKKPLMHLTLPOLQVYSGSLTLTA 300  
 QY 301 LEAKTGKLEHVENLVLMRATOLQKULTCEWGPSPKMLSLKLENKAKVSKREKPVW 360  
 DB 301 LEAKTGKLEHVENLVLMRATOLQKULTCEWGPSPKMLSLKLENKAKVSKREKPVW 360  
 QY 361 LNPBAGMOCCLSDSGQVLLBSNIXYLPWSTPVPVPCBAPBPYSCKDTHTC-----DELIG 415  
 DB 361 LNPBAGMOCCLSDSGQVLLBSNIXYLPWSTPVPVPCBAPBPYSCKDTHTC-----DELIG 415  
 QY 416 GSVFLPFPKPDOTLMSITPPEVTCVVDVSHEDPEVKVMYVDGVEVNAKTKPREBO 475  
 DB 416 GSVFLPFPKPDOTLMSITPPEVTCVVDVSHEDPEVKVMYVDGVEVNAKTKPREBO 475  
 QY 476 NSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTISAKGQPREPOVYTLPPSRD 535  
 DB 476 NSTYRVSVLTVLHODMLNGKEYKCVSNKALPAPIEKTISAKGQPREPOVYTLPPSRD 535  
 QY 536 ELTKQVSLTCLVKGFPSPDIAVEWESNQPPENNYTTPVLDSDGSFPLYSLTVDKSR 595  
 DB 536 ELTKQVSLTCLVKGFPSPDIAVEWESNQPPENNYTTPVLDSDGSFPLYSLTVDKSR 595  
 QY 541 ELTKQVSLTCLVKGFPSPDIAVEWESNQPPENNYTTPVLDSDGSFPLYSLTVDKSR 600  
 DB 541 ELTKQVSLTCLVKGFPSPDIAVEWESNQPPENNYTTPVLDSDGSFPLYSLTVDKSR 600  
 QY 596 MOGNVFCSCVWHEALHNHYTOKSLSPG 625  
 DB 601 MOGNVFCSCVWHEALHNHYTOKSLSPG 630

## RESULT 2

XX AAB19508 standard; protein; 631 AA.

XX AAB19508;

XX AAB19508; (first entry)

XX CD4-IgG1 fusion protein CH4Egamma1.

XX CD4; IgG1; human; CD4Egamma1; fusion protein; immunoglobulin; HIV; SIV;  
 KM gp120; therapy; diagnosis.  
 XX

```

OS Homo sapiens.
XX Key Location/Qualifiers
XX Protein 1..395
FT /note="CD4 extracellular region"
FT Protein 400..631
FT /note="IgG1 heavy chain"
XX US6117656-A.
XX 12-SEP-2000.
XX 07-JUN-1995; 95US-00479353.
XX 22-JAN-1988; 88US-00147351.
XX 23-JAN-1989; 89US-00289586.
XX 09-JUN-1992; 92US-00896781.
XX 12-APR-1993; 93US-00057952.
XX 04-FEB-1994; 94US-00191708.
XX (GENO ) GEN HOSPITAL CORP.
XX Seed B;
XX WPI; 2000-586558/55.
XX N-PSDB; AAA50661.
XX CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
XX SIV.
XX Example 1; Col 29-42; 39pp; English.
XX The present sequence is that of fusion protein CD4Bgamma1 comprising the
XX extracellular portion of CD4, which binds to HIV gp120, linked at its C-
XX terminus to the human IgG1 heavy chain. To obtain the fusion protein, DNA
XX encoding CD4 was linked to IgG1 DNA at the Esp site upstream of the hinge
XX region (see AAA50661). Fusion protein CD4Bgamma1 and a nucleic acid
XX encoding it are claimed. Also claimed are a vector comprising the nucleic
XX acid, and a method of producing the fusion protein in secreted form using
XX a transformed host cell. The fusion protein may further comprise a
XX therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
XX can be administered to an animal (including humans) for treatment of HIV
XX or SIV infection, and can also be used in assays for HIV or SIV, imaging
XX and tissue staining. IgG1 fusion proteins such as CD4Bgamma1 provide both
XX complement-mediated and cell-mediated immunity.
XX Sequence 631 AA:
XX
XX Query Match 94.6%; Score 3229.5; DB 3; Length 631;
XX Best Local Similarity 98.3%; Pred. No. 3e-166;
XX Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;
XX
XX 1 MNRGVPFRHLLVLIQALLPAATQGNKVVLGKGDVVELTCTASQKSIQPHMKNNOIK 60
XX 1 MNRGVPFRHLLVLIQALLPAATQGNKVVLGKGDVVELTCTASQKSIQPHMKNNOIK 60
XX 61 ILGNQGSFLTKGPKSLNDRADSRSLWQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
XX 61 ILGNQGSFLTKGPKSLNDRADSRSLWQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
XX 61 ILGNQGSFLTKGPKSLNDRADSRSLWQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120
XX 121 LVFGITANSDFHLLOQGSLLTLTLESPGSSPVOCRSRGNIOGKTLVSQLELDOSG 180
XX 121 LVFGITANSDFHLLOQGSLLTLTLESPGSSPVOCRSRGNIOGKTLVSQLELDOSG 180
XX 121 LVFGITANSDFHLLOQGSLLTLTLESPGSSPVOCRSRGNIOGKTLVSQLELDOSG 180
XX 181 TWTCITVLONQKVEFKIDIVLAPQKASSIYKKEGEVDFSPFLAFTVEKLTGSGELMW 240
XX 181 TWTCITVLONQKVEFKIDIVLAPQKASSIYKKEGEVDFSPFLAFTVEKLTGSGELMW 240
XX 181 TWTCITVLONQKVEFKIDIVLAPQKASSIYKKEGEVDFSPFLAFTVEKLTGSGELMW 240
XX 241 QAERASSSSKSWITPDLKXKEVSVKRVTPDPLQNGKPLPLHLTPQALPQYAGSGNLTIA 300
XX 241 QAERASSSSKSWITPDLKXKEVSVKRVTPDPLQNGKPLPLHLTPQALPQYAGSGNLTIA 300
XX 301 LEAKTGKLGHEVNLVVMARATQLOKNLTCEVWGPTSPKLMSLKLENKAQVSKREKPVWV 360

```

```

DB 301 LEAKTGKLGHEVNLVVMARATQLOKNLTCEVWGPTSPKLMSLKLENKAQVSKREKPVWV 360
QY 361 LNPEAGMWOCLLSGGVLLLESNIKVLPTWSTPVPCPAPEKSCDKTHTC-----PEILG 415
DB 361 LNPEAGMWOCLLSGGVLLLESNIKVLPTWSTPVHADPEEPKSCDKTHTCPCPAPELLG 420
QY 416 GPSVFLPPPKDPLMTSRPEVTCVAVDVSHDEPEVKNYVGVVHNAKTPREEROY 475
DB 421 GPSVFLPPPKDPLMTSRPEVTCVAVDVSHDEPEVKNYVGVVHNAKTPREEROY 480
QY 476 NSTRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIEKTISKAKGPREPQVYTLPSRD 535
DB 481 NSTRVVSVLTVLHQDMLNGKEYCKVSNKALPAPIEKTISKAKGPREPQVYTLPSRD 540
QY 536 ELTKNOVSLTCLVGFPPSDIAVESNGQPENNYKTPPYLSDGSFFLYSKLTVDKSR 595
DB 541 ELTKNOVSLTCLVGFPPSDIAVESNGQPENNYKTPPYLSDGSFFLYSKLTVDKSR 600
QY 596 WQGNVFSQSYMHEALNHYTQKSLSPG 625
DB 601 WQGNVFSQSYMHEALNHYTQKSLSPG 630

RESULT 3
AA51079
ID AA51079 standard; protein; 631 AA.
XX
XX AA51079;
XX
XX 23-MAR-2000 (first entry)
XX
XX Human fusion protein CD4Bgamma1.
XX
XX Fusion protein; human; CD4; IgG1; immunoglobulin; gp120;
XX anti-human immunodeficiency virus; CD4Bgamma1.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX US6004781-A.
XX
XX 21-DEC-1999.
XX
XX 04-FEB-1994; 94US-00191708.
XX
XX 22-JAN-1988; 88US-00147351.
XX 23-JAN-1989; 89US-00289586.
XX 09-JUN-1992; 92US-00896781.
XX 12-APR-1993; 93US-00057952.
XX
XX (GENO ) GEN HOSPITAL CORP.
XX
XX Seed B;
XX WPI; 2000-085792/07.
XX N-PSDB; AA244062.
XX
XX Fusion protein useful for the treatment of human immunodeficiency virus.
XX
XX Example 1; Col 29-42; 39pp; English.
XX
XX This invention describes a novel nucleic acid (I) encoding a fusion
XX protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
XX and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
XX chain (III). The products of the invention have anti-human
XX immunodeficiency virus (HIV) activity and are capable of binding to
XX gp120. The fusion protein is useful for treating human immunodeficiency
XX virus (HIV) or simian immunodeficiency virus (SIV). This sequence
XX represents the fusion protein CD4Bgamma1 which is constructed from CD4
XX linked to human IgG1 upstream of the hinge region
XX
XX Sequence 631 AA:

```

Query Match 94.6%; Score 3229.5; DB 3; Length 631;  
Best Local Similarity 98.3%; Pred. No. 3e-166;  
Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;

QY 1 MNRGVPRHLLVLTQALLPAATQGNKVTLGKKGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVPRHLLVLTQALLPAATQGNKVTLGKKGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGSPSKINDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
DB 61 ILGNQGSFLTKGSPSKINDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
QY 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLOQOKKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240  
DB 181 TWICTVLOQOKKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240  
QY 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300  
DB 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300  
QY 301 LEAKTGKLEHQBVLVVMRATQLOKNTLCEVWGPTSPKMLSLLENKAKVSKREKRVWY 360  
DB 301 LEAKTGKLEHQBVLVVMRATQLOKNTLCEVWGPTSPKMLSLLENKAKVSKREKRVWY 360  
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPAPAPKSCDTHTC-----PELLG 415  
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPAPAPKSCDTHTC-----PELLG 415  
QY 416 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475  
DB 416 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475  
QY 421 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 480  
DB 421 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 480  
QY 476 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 535  
DB 476 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 535  
QY 481 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 540  
DB 481 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 540  
QY 536 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 595  
DB 536 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 595  
QY 541 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 600  
DB 541 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 600  
QY 596 WQGNVFSQSVMEHALLHNHYTQKSLSLSPG 625  
DB 601 WQGNVFSQSVMEHALLHNHYTQKSLSLSPG 630

RESULT 4  
AAV59169  
ID AAV59169 standard; protein; 631 AA.  
XX  
AC AAV59169;  
XX  
DT 14-MAR-2000 (first entry)  
XX  
DE CD4-Ig fusion protein CD4Egammal.  
XX  
KM HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
XX secreted protein; HIV infection; medicament.  
XX  
XX Synthetic.  
XX OS Homo sapiens.  
XX PN CAL340741-C.  
XX  
XX 14-SEP-1999.  
XX PD  
XX PF 20-JAN-1989; 89CA-00588749.  
XX PR 20-JAN-1989; 89CA-00588749.

XX (GENE) GEN HOSPITAL CORP.  
PA Seed B;  
XX  
XX MPI; 2000-063015/06.  
DR N-PSDB; A4248202.  
PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
XX the treatment of HIV or simian immunodeficiency virus infections.  
XX  
XX Example 1; Page 37-46; 89pp; English.  
XX  
CC The invention provides a fusion gene encoding a fusion protein that  
CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
CC the variable region has been replaced with the DNA sequence which encodes  
CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
CC capable of being secreted. The fusion proteins are useful for treating  
CC HIV or SIV infections in animals, preferably humans. They are also useful  
CC for producing medicaments which can be used for treating HIV or SIV  
CC infections in humans. The present sequence represents the fusion protein  
CC CD4Egammal where the CD4 is linked to human IgG1 at the Bsp site upstream  
CC of the hinge region  
XX  
SQ Sequence 631 AA;  
XX  
Query Match 94.6%; Score 3229.5; DB 3; Length 631;  
Best Local Similarity 98.3%; Pred. No. 3e-166;  
Matches 619; Conservative 0; Mismatches 6; Indels 5; Gaps 1;

QY 1 MNRGVPRHLLVLTQALLPAATQGNKVTLGKKGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVPRHLLVLTQALLPAATQGNKVTLGKKGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGSPSKINDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
DB 61 ILGNQGSFLTKGSPSKINDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
QY 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLOQGSJLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLOQOKKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240  
DB 181 TWICTVLOQOKKVEFKIDIVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240  
QY 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300  
DB 241 QAERASSSKSWITFDLKNKEVSVKRVYQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300  
QY 301 LEAKTGKLEHQBVLVVMRATQLOKNTLCEVWGPTSPKMLSLLENKAKVSKREKRVWY 360  
DB 301 LEAKTGKLEHQBVLVVMRATQLOKNTLCEVWGPTSPKMLSLLENKAKVSKREKRVWY 360  
QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPAPAPKSCDTHTC-----PELLG 415  
DB 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPAPAPKSCDTHTC-----PELLG 415  
QY 416 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475  
DB 416 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 475  
QY 421 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 480  
DB 421 GPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGEVHNAAKTKPREEOY 480  
QY 476 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 535  
DB 476 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 535  
QY 481 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 540  
DB 481 NSTYRVVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGQPREQVYTTLPSPRD 540  
QY 536 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 595  
DB 536 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 595  
QY 541 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 600  
DB 541 ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSR 600  
QY 596 WQGNVFSQSVMEHALLHNHYTQKSLSLSPG 625

601 MOGNVFCSCWHEALHNHYTKSLSTLSFG 630

```
Dh      |||
RESULT 5
AAP93008 AAP93008 standard; protein; 729 AA.
XX ID     AAP93008 standard; protein; 729 AA.
AC       AAP93008;
XX DT     25-MAR-2003 (revised)
DT       02-NOV-1992 (first entry)
XX XX
DE Genetic construct which encodes CD4 linked to human IgG1 at the Hind3 site upstream of the CH1 region (fusion protein CD4H-gamma-1).
KW Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy; diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.
XX KM Homo sapiens.
OS
CS EPJ25262-A.
PN
PD 26-JUL-1989.
PF 20-JAN-1989; 89EP-00100913.
PR 22-JAN-1988; 88US-00147351.
PP XA (GEHO ) GEN HOSPITAL CORP.
PA Seed B;
PI
XX DR WPI; 1989-214472/30.
DR N-PADB; AAN90356.
XX PT Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV infections or detecting HIV or SIV in sample.
XX PS Example; Table 1, Page 12-23; 68pp; English.
XX CC The fusion protein genes of the invention pref. comprise cDNA sequences which encode CD4 or a fragment which binds gp120 ligated to an expression plasmid which encodes an antibody in which the variable region of the gene has been deleted (see WO87-02671). The CD4 portion of the fusion protein may comprise the complete CD4 sequence, the 370 AA extracellular region and the membrane spanning domain, or the extracellular region. The heavy chain is pref. from IgM, IgG1 or IgG3. The following are specifically claimed: fusion proteins CD4H-gamma-1, CD4mu, CD4pm, CD4E-gamma1, and CD4mmu (No. 67608), PCDAF-gamma (No. 67609) and PCDAE-gamma-1 (No. 67610). The plasmid containing (PCDAH-gamma-1) has been deposited in E. coli (MC1061/B3) at the ATCC under accession number 67611. (Updated on 25-MAR-2003 to correct PA field.)
XX SC Sequence 729 AA;
```

```
Query Match          93.0%; Score 3176.5; DB 1; Length 729;
Best Local Similarity 85.3%; Pred. No. 2.5e+163;
Matches   621; Conservative    0; Mismatches    4; Indels 103; Gaps    3
```

```
OY 1 MRGRVPFHLLLVLTQLLIPATQGNKVKLGKKGTVELTCTASOGKSIOFMKNNSNQIK 60
Db 1 MNRGVFPHLLLVLTLQALLIPAATQGKVVLGGKGDTVETLTCAOQSIOFMKNNSNQT 60
YY ILNGSGFLTGPSRLNDRAISRILMOAGNPPLIIKLKIEDSDTYICEVDOKEEVGL 120
Dd 61 I LGNGSFLTGPSTLNDRADSRIRLMQGNFPPLI IKLKIEDSDTYICEVDOKEEVGL 120
OY 1 L VFGITANSPTHTLLOGOSTLTTLLESAPPGSSPEVCQRSPRCKNI QCGKTLSVSQLEIDBSG 180
DY 121 LV FGI TANS DTH L LOG OS TL TL LES AP PG SS PE VC QR SP RK GN IO CG K TS VS Q LE IB SG 180
DB 121 LV FG IT AN SD TH LO GO ST LT TL ES PP GS SPE VC QS PR GK NI OG CK TS VS Q LE IB DS G 180
OY 181 TWTCVTLO NKKVEFKIDI VLVAFOKASSIVYKKEGOVERSFPLAFIVEKLTGSGETLWW 240
```

Accession	Protein	Length	Source
Db	161 TWTCVLQNKVVERKIDIVLAFQKASIVYKKEGQVEFSPLAFVTEKLTGSGELMM	240	QAEARASSSKSWITTFDLKKNKESVSKVKTQDPKLOMGKCLPLHLTLPOLPOLQVAGSGMLTTLA
Qy	241 QAEARASSSKSWITTFDLKKNKESVSKVKTQDPKLOMGKCLPLHLTLPOLPOLQVAGSGMLTTLA	300	241 QAEARASSSKSWITTFDLKKNKESVSKVKTQDPKLOMGKCLPLHLTLPOLPOLQVAGSGMLTTLA
Db	301 LEAKTGKIHQEVNVLVMMRATQLOKNLTCEVMGPTSPKJMLSLKLENKEAKVSKREKPVWV	360	301 LEAKTGKIHQEVNVLVMMRATQLOKNLTCEVMGPTSPKJMLSLKLENKEAKVSKREKPVWV
Qy	301 LEAKTGKIHQEVNVLVMMRATQLOKNLTCEVMGPTSPKJMLSLKLENKEAKVSKREKPVWV	360	301 LEAKTGKIHQEVNVLVMMRATQLOKNLTCEVMGPTSPKJMLSLKLENKEAKVSKREKPVWV
Db	361 LNPEAGMWCCLISDSGQVLLBSNITVLPWMSGTPVC-----	399	361 LNPEAGMWCCLISDSGQVLLBSNITVLPWMSGTPVC-----
Qy	397 -----PAP-----	399	397 -----PAP-----
Db	421 TTAALGCLVKDYFRFEVTVVSWNSGALITSGVHTPRVAVLQSSGLXLSLSVTVTPSSSLGTQTY	480	421 TTAALGCLVKDYFRFEVTVVSWNSGALITSGVHTPRVAVLQSSGLXLSLSVTVTPSSSLGTQTY
Qy	400 -----EPKSCDKHTHTC-----PILGGPSVFLPFPKPKDTLMISRTPE	437	400 -----EPKSCDKHTHTC-----PILGGPSVFLPFPKPKDTLMISRTPE
Db	481 ICNVNHHKPSNTKVDKKEVPEKSCDKHTHTCPCRPAPELLGGPSVFLPFPKPKDTLMISRTPE	540	481 ICNVNHHKPSNTKVDKKEVPEKSCDKHTHTCPCRPAPELLGGPSVFLPFPKPKDTLMISRTPE
Qy	438 VTCVVVDVSHEDPEVKFNWYVDGVEVNAKATKPREQVNSTTRVYSVLTLYLQDYLNGKE	497	438 VTCVVVDVSHEDPEVKFNWYVDGVEVNAKATKPREQVNSTTRVYSVLTLYLQDYLNGKE
Db	541 VTCVVVDVSHEDPEVKFNWYVDGVEVNAKATKPREQVNSTTRVYSVLTLYLQDYLNGKE	600	541 VTCVVVDVSHEDPEVKFNWYVDGVEVNAKATKPREQVNSTTRVYSVLTLYLQDYLNGKE
Qy	498 YKCKVSNKALPAPIEKTTISKAKGQREPOVYTLPPSRDELTKGNVSLTCLVKGFPYSDIA	557	498 YKCKVSNKALPAPIEKTTISKAKGQREPOVYTLPPSRDELTKGNVSLTCLVKGFPYSDIA
Db	601 YKCKVSNKALPAPIEKTTISKAKGQREPOVYTLPPSRDELTKGNVSLTCLVKGFPYSDIA	660	601 YKCKVSNKALPAPIEKTTISKAKGQREPOVYTLPPSRDELTKGNVSLTCLVKGFPYSDIA
Qy	558 VEWESNGQRENNYKTTTPVLDSGDSFFLYSKLTVDKSRKQDGNVSCSVMEHALNNHYTQ	617	558 VEWESNGQRENNYKTTTPVLDSGDSFFLYSKLTVDKSRKQDGNVSCSVMEHALNNHYTQ
Db	661 VEWESNGQRENNYKTTTPVLDSGDSFFLYSKLTVDKSRKQDGNVSCSVMEHALNNHYTQ	720	661 VEWESNGQRENNYKTTTPVLDSGDSFFLYSKLTVDKSRKQDGNVSCSVMEHALNNHYTQ
Qy	618 KSLSLSPG 625		618 KSLSLSPG 625
Db	721 KSLSLSPG 728		721 KSLSLSPG 728

PA (GEHO ) GEN HOSPITAL CORP.  
 XX Seed B;  
 PI  
 XX WPI; 2000-586558/55.  
 DR N-PSDB; AAA50660.  
 XX  
 PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or  
 PT SIV.  
 XX  
 XX Example 1; Col 13-30; 39pp; English.  
 XX  
 CC The present sequence is that of fusion protein CD4Hgamma1 comprising the  
 CC extracellular portion of CD4, which binds to HIV gp120, linked at its C-  
 CC terminus to the human IgG1 heavy chain. To obtain the fusion protein, DNA  
 CC encoding CD4 was linked to IgG1 DNA at the Hind3 site upstream of the CH1  
 CC region (see AAA50660). Fusion protein CD4Hgamma1 and a nucleic acid  
 CC encoding it are claimed. Also claimed are a vector comprising the nucleic  
 CC acid, and a method of producing the fusion protein in secreted form using  
 CC a transformed host cell. The fusion protein may further comprise a  
 CC therapeutic agent, radiolabel or NMR imaging agent. The fusion protein  
 CC can be administered to an animal (including humans) for treatment of HIV  
 CC or SIV infection, and can also be used in assays for HIV or SIV, imaging  
 CC and tissue stains. IgG1 fusion proteins such as CD4Hgamma1 provide both  
 CC complement-mediated and cell-mediated immunity  
 XX  
 SQ Sequence 729 AA;  
 Query Match 93.0%; Score 3176.5; DB 3; Length 729;  
 Best Local Similarity 85.3%; Pred. No. 2.5e-163;  
 Matches 621; Conservative 0; Mismatches 4; Indels 103; Gaps 3;  
 QY 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNOIK 60  
 QY 61 ILNGQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120  
 DB 61 ILNGQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120  
 QY 121 LVFGLTANSDTHLLQGSLLTLLTLESPPGSSPSVQCRSPRKNIOGCKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGSLLTLLTLESPPGSSPSVQCRSPRKNIOGCKTLVSQLELDQSG 180  
 QY 181 TWTCVTIQQNKQKVEFKIDIVLAFQKASSIYKKESQVEPSFPLATVEKLTGSGELMW 240  
 DB 181 TWTCVTIQQNKQKVEFKIDIVLAFQKASSIYKKESQVEPSFPLATVEKLTGSGELMW 240  
 QY 241 QAERASSKSMITFDLKNKESVSRVQDPRKIQMGKLLPHLTLPOLPYAGSGNLTLA 300  
 DB 241 QAERASSKSMITFDLKNKESVSRVQDPRKIQMGKLLPHLTLPOLPYAGSGNLTLA 300  
 QY 301 LEAKTGLHQBVLNVMKATQLOKNTLCEVWGPTSPKMLSLKLENKAVSKREKPVWV 360  
 DB 301 LEAKTGLHQBVLNVMKATQLOKNTLCEVWGPTSPKMLSLKLENKAVSKREKPVWV 360  
 QY 361 LNPBAGWQCLISGQVLLLESNIKVLPTWSTPVPFC----- 396  
 DB 361 LNPBAGWQCLISGQVLLLESNIKVLPTWSTPVPFC----- 396  
 QY 397 -----PAP----- 399  
 DB 421 TAAIGCLVSTFPEPVTVMNSGALITSGVHTFPAVLQSSGLYSLSAVTVPPSSSLGTQTY 480  
 QY 400 -----EPKSCDKTHTC-----PELLGSPVFLFPKPKDTLMISRTPE 437  
 DB 481 ICNVNHNKPSMTKVDKVEPKSCDKTHTCPCPAPELLGSPVFLFPKPKDTLMISRTPE 540  
 QY 438 VTCVAVDVSHEDPEVKKNMYYDGVFNNAKTKPREEDYNSTYRVVSVLYLVHODMLNGKE 497  
 DB 541 VTCVAVDVSHEDPEVKKNMYYDGVFNNAKTKPREEDYNSTYRVVSVLYLVHODMLNGKE 600  
 QY 498 YKCKVSKALPAPLEKTIKAKQGPREPQVYTLPPSDELTKQGVSLTCLVKQGYPSDIA 557

DB 601 YKCKVSKALPAPLEKTIKAKQGPREPQVYTLPPSDELTKQGVSLTCLVKQGYPSDIA 660  
 QY 558 VEMESNGQENNYKTTTPVLDSDGSEFLYSKLTVDKSRMOQGNVFCSVMAEALNNHYTQ 617  
 DB 661 VEMESNGQENNYKTTTPVLDSDGSEFLYSKLTVDKSRMOQGNVFCSVMAEALNNHYTQ 720  
 QY 618 KSLSLSPG 625  
 DB 721 KSLSLSPG 728  
 RESULT 7  
 AA59168  
 ID AA59168 standard; protein; 729 AA.  
 AC AA59168;  
 XX 14-MAR-2000 (first entry)  
 DT  
 XX CD4-Ig fusion protein CD4Hgamma1.  
 DE  
 XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
 KM secreted protein; SIV infection; medicament.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 PN CA1340741-C.  
 PD 14-SEP-1999.  
 PF 20-JAN-1989; 89CA-00586749.  
 PR 20-JAN-1989; 89CA-00586749.  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX Seed B;  
 PI  
 XX WPI; 2000-063015/06.  
 DR N-PSDB; AA428201.  
 XX  
 PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
 PT the treatment of HIV or simian immunodeficiency virus infections.  
 XX  
 XX Example 1; Page 25-36; 89pp; English.  
 CC The invention provides a fusion gene encoding a fusion protein that  
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 CC the variable region has been replaced with the DNA sequence which encodes  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the fusion protein  
 CC CD4Hgamma1 where the CD4 is linked to human IgG1 at the Hind3 site  
 CC upstream of the CH1 region  
 XX  
 SQ Sequence 729 AA;  
 Query Match 93.0%; Score 3176.5; DB 3; Length 729;  
 Best Local Similarity 85.3%; Pred. No. 2.5e-163;  
 Matches 621; Conservative 0; Mismatches 4; Indels 103; Gaps 3;  
 QY 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPFRHLVLVQLALLPAATQGNKVLGKKGDVVELTCTASOKKSIOFHMKNNOIK 60  
 QY 61 ILNGQGSFLTKGPKSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEVOL 120

```

Db      61 ILGNQSFLLTKGPKANDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEYOL 120
Qy      121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
Db      181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
Qy      241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQMGKKLPLHLTLPLQALPYAGSGNLTLA 300
Db      241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQMGKKLPLHLTLPLQALPYAGSGNLTLA 300
Qy      301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
Db      301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
Qy      361 LNPEAGMWQCILSDSGQVLLSNIKVLPWTMSTPVPC----- 396
Db      361 LNPEAGMWQCILSDSGQVLLSNIKVLPWTMSTPVHADPEASTKGSVFPPLASSKSTSGG 420
Qy      397 -----PAP----- 399
Db      421 TAAIGCLVKDYFPEPVTVSNMNSGALTSQHTFPAVLQSSGLYSLSVTVTPSSSLGTQTY 480
Qy      400 -----EPKSCDKTHTC-----PELLGSPSVLPFPKPKDTLMISRTPE 437
Db      481 ICNVNHRKSNTRKVDKVEPKSCDKTHTCPCCPAPPELLGSPSVLPFPKPKDTLMISRTPE 540
Qy      438 VTCVVDVSHEDPEVKFMVYVDGVEVHNAAKTPREEQYNSTYRVVSVLTVLHODMLNGKE 497
Db      541 VTCVVDVSHEDPEVKFMVYVDGVEVHNAAKTPREEQYNSTYRVVSVLTVLHODMLNGKE 600
Qy      498 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNQVSLTCLVKGYFSPDIA 557
Db      601 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNQVSLTCLVKGYFSPDIA 660
Qy      558 VEMESNQGPENNYKTPPVLDSDGSFLYSKLTVDKSRMWOQGNVFSCSVMHEALHNHYTQ 617
Db      661 VEMESNQGPENNYKTPPVLDSDGSFLYSKLTVDKSRMWOQGNVFSCSVMHEALHNHYTQ 720
Qy      618 KSLSLSPG 625
Db      721 KSLSLSPG 728

RESULT 8
ID AAYS1078 standard; protein, 729 AA.
AC AAYS1078;
XX
XX 23-MAR-2000 (first entry)
XX
XX Human fusion protein CD4H-1.
XX
XX Fusion protein; human, CD4, IgG1, immunoglobulin; gp120;
XX anti-human immunodeficiency virus; CD4H-1.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX US6004781-A.
XX
XX 21-DEC-1999.
XX
XX 04-FEB-1994; 94US-00191708.
XX
XX 22-JAN-1988; 88US-00147351.
XX 23-JAN-1989; 89US-00299596.
XX 09-JUN-1992; 92US-00896781.
XX 12-APR-1993; 93US-00057952.

```

```

XX      XX (GENO ) GEN HOSPITAL CORP.
PA      XX
XX      XX Seed B;
PI      XX
XX      XX WPI; 2000-085792/07.
DR      XX
XX      XX N-PSDB; AAZ44061.
PT      XX
XX      XX Fusion protein useful for the treatment of human immunodeficiency virus.
PS      XX
XX      XX Example 1; Col 15-30; 39pp; English.
CC      XX
CC      XX This invention describes a novel nucleic acid (I) encoding a fusion
CC      XX protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC      XX and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC      XX chain (III). The products of the invention have anti-human
CC      XX immunodeficiency virus (HIV) activity and are capable of binding to
CC      XX gp120. The fusion protein is useful for treating human immunodeficiency
CC      XX virus (HIV) or simian immunodeficiency virus (SIV). This sequence
CC      XX represents the fusion protein CD4H-1 which is constructed from CD4 linked
CC      XX to human IgG1 upstream of the CH1 region
XX      XX
XX      XX Sequence 729 AA:
S0
Query Match      92.8%; Score 3166.5; DB 3; Length 729;
Best Local Similarity 85.2%; Pred. No. 8.8e-163;
Matches 620; Conservative 0; Mismatches 5; Indels 103; Gaps 3;

Qy      1 MNRGVPFRHLVLVQLALPAPATQGNKRVVLGKKGDTVELTCTASQKKSIOFHMNSNQIK 60
Db      1 MNRGVPFRHLVLVQLALPAPATQGNKRVVLGKKGDTVELTCTASQKKSIOFHMNSNQIK 60
Qy      61 ILGNQSFLLTKGPKANDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEYOL 120
Db      61 ILGNQSFLLTKGPKANDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEYOL 120
Qy      121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
Db      181 TWCTVLQONQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFTVEKLTGSGELMW 240
Qy      241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQMGKKLPLHLTLPLQALPYAGSGNLTLA 300
Db      241 QAEARASSKSWITFDLKNKEVSVKRVYODPKLQMGKKLPLHLTLPLQALPYAGSGNLTLA 300
Qy      301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
Db      301 LEAKTGKHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKEAVSKREKPVWV 360
Qy      361 LNPEAGMWQCILSDSGQVLLSNIKVLPWTMSTPVPC----- 396
Db      361 LNPEAGMWQCILSDSGQVLLSNIKVLPWTMSTPVHADPEASTKGSVFPPLASSKSTSGG 420
Qy      397 -----PAP----- 399
Db      421 TAAIGCLVKDYFPEPVTVSNMNSGALTSQHTFPAVLQSSGLYSLSVTVTPSSSLGTQTY 480
Qy      400 -----EPKSCDKTHTC-----PELLGSPSVLPFPKPKDTLMISRTPE 437
Db      481 ICNVNHRKSNTRKVDKVEPKSCDKTHTCPCCPAPPELLGSPSVLPFPKPKDTLMISRTPE 540
Qy      438 VTCVVDVSHEDPEVKFMVYVDGVEVHNAAKTPREEQYNSTYRVVSVLTVLHODMLNGKE 497
Db      541 VTCVVDVSHEDPEVKFMVYVDGVEVHNAAKTPREEQYNSTYRVVSVLTVLHODMLNGKE 600
Qy      498 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNQVSLTCLVKGYFSPDIA 557
Db      601 YKCVSNKALPAPIEKTISKAKGQPREPQVYTLPPSDELTKNQVSLTCLVKGYFSPDIA 660
Qy      558 VEMESNQGPENNYKTPPVLDSDGSFLYSKLTVDKSRMWOQGNVFSCSVMHEALHNHYTQ 617

```



```
Db      661 VEEESNGQPENNKTTTPVLDSDGSFFLYSKLTVDSRMQOGNVFSCSVNHEALHNHYQ 720
QY      618 KSLSLSPG 625
Db      721 KSLSLSPG 728

RESULT 9
AA51082
ID      AAY51082 standard; protein; 616 AA.
XX
AC      AAY51082;
XX
DT      23-MAR-2000 (first entry)
XX
DE      Human fusion protein CD4Bgammal.
XX
KM      Fusion protein; human; CD4; IgG1; immunoglobulin; gp120;
KM      anti-human immunodeficiency virus; CD4Bgammal.
XX
OS      Homo sapiens.
OS      Synthetic.
XX
PN      US6004781-A.
XX
PD      21-DEC-1999.
XX
PF      04-FEB-1994; 94US-00191708.
XX
PR      22-JAN-1988; 88US-00147351.
PR      23-JAN-1989; 89US-00295956.
PR      09-JUN-1982; 92US-008956781.
PR      12-APR-1993; 93US-00057952.
XX
PA      (GEOH ) GEN HOSPITAL CORP.
XX
PI      Seed B;
XX
DR      WPI; 2000-085792/07.
XX
DR      N-PSDB; AAZ44065.
XX
PT      Fusion protein useful for the treatment of human immunodeficiency virus.
XX
PS      Example 1; Col 59-70; 39pp; English.
XX
CC      This invention describes a novel nucleic acid (I) encoding a fusion
CC      protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC      and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC      chain (III). The products of the invention have anti-human
CC      immunodeficiency virus (HIV) activity and are capable of binding to
CC      gp120. The fusion protein is useful for treating human immunodeficiency
CC      virus (HIV) or simian immunodeficiency virus (SIV). This sequence
CC      represents the fusion protein CD4Bgammal which is constructed from CD4
CC      linked to human IgG1 upstream of the hinge region
XX
SQ      Sequence 616 AA;

Query Match      92.6%; Score 3163; DB 3; Length 616;
Best Local Similarity 97.6%; Pred. No. 1,1e-162;
Matches 610; Conservative 0; Mismatches 5; Indels 10; Gaps 2;
```

```
QY      181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGQVEFSPPLAFVTEKLTGSGELWM 240
Db      181 TWCTVLONOKKVEFKIDIVLAFQKASSIYKKEGQVEFSPPLAFVTEKLTGSGELWM 240
QY      241 QAERASSSKSWITFDLKNKEVSVKRVTQDPKLOMGKULPLHLTLPOLPOYAGSGNLTLA 300
Db      241 QAERASSSKSWITFDLKNKEVSVKRVTQDPKLOMGKULPLHLTLPOLPOYAGSGNLTLA 300
QY      301 LEAKTGKLEHVEVNLVVMRATQLOKULTCVWGPSTSPKMLSLKENKAKVSKREPVW 360
Db      301 LEAKTGKLEHVEVNLVVMRATQLOKULTCVWGPSTSPKMLSLKENKAKVSKREPVW 360
QY      361 LNPEAGMOCCLSDSGQVLEESNIKYLPTWSTPVPAPREPSCDKTHCPELLGSPSYF 420
Db      361 LNPEAGMOCCLSDSGQVLEESNIKYLPTWSTPVPAD-PE-----APELLGSPSYF 410
QY      421 LEPPKPKDTLMSRTPPEVTCVVVDVSHEDPEVKFMWYDGVENHAKTRPREQYNSTYR 480
Db      411 LEPPKPKDTLMSRTPPEVTCVVVDVSHEDPEVKFMWYDGVENHAKTRPREQYNSTYR 470
QY      481 VVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKV 540
Db      471 VVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKV 530
QY      541 QVSLTCLVKGFPSPDIAVESNGQPENNKTTTPVLDSDGSFFLYSKLTVDSRMQOGN 600
Db      531 QVSLTCLVKGFPSPDIAVESNGQPENNKTTTPVLDSDGSFFLYSKLTVDSRMQOGN 590
QY      601 VFSCSVMEALHNHYTQKSLSLSPG 625
Db      591 VFSCSVMEALHNHYTQKSLSLSPG 615

RESULT 10
AA59172
ID      AAY59172 standard; protein; 616 AA.
XX
AC      AAY59172;
XX
DT      14-MAR-2000 (first entry)
XX
DE      CD4-Ig fusion protein CD4Bgammal.
XX
KM      HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
KM      secreted protein; SIV infection; medicament.
XX
OS      Synthetic.
OS      Homo sapiens.
XX
PN      CA1340741-C.
XX
PD      14-SEP-1999.
XX
PF      20-JAN-1989; 89CA-00588749.
XX
PR      20-JAN-1989; 89CA-00588749.
XX
PA      (GEOH ) GEN HOSPITAL CORP.
XX
PI      Seed B;
XX
DR      WPI; 2000-063015/06.
XX
DR      N-PSDB; AAZ48205.
XX
PT      New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
PT      the treatment of HIV or simian immunodeficiency virus infections.
XX
PS      Example 1; Page 61-68; 89pp; English.
XX
CC      The invention provides a fusion gene encoding a fusion protein that
CC      comprises an extracellular CD4 DNA sequence or its fragment which binds
CC      to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
```

CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 CC the variable region has been replaced with the DNA sequence which encodes  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the fusion protein  
 CC CD4Bgamma where the CD4 is linked to human IgG1 at the BamI site  
 CC downstream from the hinge region

XX Sequence 616 AA:

Query Match 92.6%; Score 3163; DB 3; Length 616;

Best Local Similarity 97.6%; Pred. No. 1,1e-162;

Matches 610; Conservative 0; Mismatches 5; Indels 10; Gaps 2;

QY 1 MNRGVPRHLLLVQLALLPAATQGNKRVLGKGGDTVELCTASOKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPRHLLLVQLALLPAATQGNKRVLGKGGDTVELCTASOKKSIOFHMKNNOIK 60  
 QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIVYKKEGEVFSPLAFTVEKLTGSGELMW 240  
 DB 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIVYKKEGEVFSPLAFTVEKLTGSGELMW 240  
 QY 241 QAERASSSKSMITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPLQALPOYAGSGNLTIA 300  
 DB 241 QAERASSSKSMITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPLQALPOYAGSGNLTIA 300  
 QY 301 LEAKTGKLEHENVLVNRRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360  
 DB 301 LEAKTGKLEHENVLVNRRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360  
 QY 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLEPTWSTPVPCAPBEPKSCDKHTCPBELLGSPVF 420  
 DB 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLEPTWSTPVPCAPBEPKSCDKHTCPBELLGSPVF 420  
 QY 421 LFPKPKDPTLMISTRTPEVTCVVVDVSHEDPEVKFNNYVDDGEVHNAAKTKRREDOYNSTYR 480  
 DB 421 LFPKPKDPTLMISTRTPEVTCVVVDVSHEDPEVKFNNYVDDGEVHNAAKTKRREDOYNSTYR 480  
 QY 481 VSVLTVLHODMLNGKEYCKKCNKALPAPIETKTSKAKGQPREPOVYTLPPSRDELTKN 540  
 DB 481 VSVLTVLHODMLNGKEYCKKCNKALPAPIETKTSKAKGQPREPOVYTLPPSRDELTKN 540  
 QY 541 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTPRPVLDSDGSPFLYSKLTVDKSRMOQGN 600  
 DB 541 QVSLTCLVKGFPYPSDIAVEMESNGQPENNYKTPRPVLDSDGSPFLYSKLTVDKSRMOQGN 600  
 QY 601 VFSCSVHGEALHNHYTQKSISLSPG 625  
 DB 601 VFSCSVHGEALHNHYTQKSISLSPG 625  
 QY 591 VFSCSVHGEALHNHYTQKSISLSPG 615  
 DB 591 VFSCSVHGEALHNHYTQKSISLSPG 615

RESULT 11  
 ID AAP93012 standard; protein; 614 AA.  
 AC AAP93012;  
 XX 25-MAR-2003 (revised)  
 DT 03-AUG-1992 (first entry)  
 XX Genetic construct which encodes CD4 linked to human IgG1 at the BamI site  
 DE downstream from the hinge region (fusion protein CD4Bgamma).

KW Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy;  
 KW diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.

OS Homo sapiens.

PN EP325262-A.

XX 26-JUL-1989.

PP 20-JAN-1989; 89EP-00100913.

PR 22-JAN-1988; 88US-00147351.

PA (GENO) GEN HOSPITAL CORP.

XX Seed B;

DR WPI; 1989-214472/30.

DR N-PSDB; AAN90360.

PT Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV  
 PT infections or detecting HIV or SIV in sample.

PS Example; Table 5, Page 48-55; 68pp; English.

CC The fusion protein genes of the invention pref. comprises cDNA sequences  
 CC which encode CD4 or a fragment which binds gp120 ligated to an expression  
 CC plasmid which encodes an antibody in which the variable region of the  
 CC gene has been deleted (see WO87-02671). The CD4 portion of the fusion  
 CC protein may comprise the complete CD4 sequence, the 370 AA extracellular  
 CC region and the membrane spanning domain, or the extracellular region. The  
 CC Ig heavy chain is pref. from IgM, IgG1 or IgG3. The following are  
 CC specifically claimed: fusion proteins CD4Bgamma, CD4mu, CD4Bmu,  
 CC CD4Bdelta, and CD4mu (No. 67608), CD4Bgamma (No. 67609) and  
 CC CD4Bdelta (No. 67610). (Updated on 25-MAR-2003 to correct PA field.)

SO Sequence 614 AA:

Query Match 91.7%; Score 3132; DB 1; Length 614;

Best Local Similarity 97.0%; Pred. No. 5.3e-161;

Matches 606; Conservative 0; Mismatches 7; Indels 12; Gaps 3;

QY 1 MNRGVPRHLLLVQLALLPAATQGNKRVLGKGGDTVELCTASOKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPRHLLLVQLALLPAATQGNKRVLGKGGDTVELCTASOKKSIOFHMKNNOIK 60  
 QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIVYKKEGEVFSPLAFTVEKLTGSGELMW 240  
 DB 181 TWCTCTVLONQKKEFKIDIVLAFQKASSIVYKKEGEVFSPLAFTVEKLTGSGELMW 240  
 QY 241 QAERASSSKSMITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPLQALPOYAGSGNLTIA 300  
 DB 241 QAERASSSKSMITFDLKNKEVSVKRVYQDPKLOMGKPLHLTLPLQALPOYAGSGNLTIA 300  
 QY 301 LEAKTGKLEHENVLVNRRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360  
 DB 301 LEAKTGKLEHENVLVNRRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360  
 QY 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLEPTWSTPVPCAPBEPKSCDKHTCPBELLGSPVF 420  
 DB 361 LNPBAGMWQCLLSDSGQVLLSNNIKVLEPTWSTPVPCAPBEPKSCDKHTCPBELLGSPVF 420  
 QY 421 LFPKPKDPTLMISTRTPEVTCVVVDVSHEDPEVKFNNYVDDGEVHNAAKTKRREDOYNSTYR 480  
 DB 421 LFPKPKDPTLMISTRTPEVTCVVVDVSHEDPEVKFNNYVDDGEVHNAAKTKRREDOYNSTYR 480  
 QY 411 LFPKPKDPTLMISTRTPEVTCVVVDVSHEDPEVKFNNYVDDGEVHNAAKTKRREDOYNSTYR 470  
 DB 411 LFPKPKDPTLMISTRTPEVTCVVVDVSHEDPEVKFNNYVDDGEVHNAAKTKRREDOYNSTYR 470

```

QY 481 VSVLTIVLHODWLNCKEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 540
DB 471 VSVLTIVLHODWLNCKEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 528
QY 541 QVSLTCLVKGFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDSKRMQGN 600
DB 529 QVSLTCLVKGFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDSKRMQGN 588
QY 601 VFSCSVMEHALHNHYTKSLSLSPG 625
DB 589 VFSCSVMEHALHNHYTKSLSLSPG 613

RESULT 12
AAB19511
ID AAB19511 standard; protein; 616 AA.
XX
AC AAB19511;
XX
DT 09-JAN-2001 (first entry)
XX
DE CD4-IgG1 fusion protein CH4Bgammal.
XX
KM CD4; IgG1; human; CD4Bgammal; fusion protein; immunoglobulin; HIV; SIV;
KM gp120; therapy; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..395
FT Protein /note="CD4 extracellular region"
FT Protein 400..616
FT Protein /note="IgG1 heavy chain"
XX
XX US6117656-A.
XX
PD 12-SEP-2000.
XX
PF 07-JUN-1995; 95US-00479353.
XX
PR 22-JAN-1988; 88US-00147351.
PR 23-JAN-1989; 89US-00299596.
PR 09-JUN-1992; 92US-0086781.
PR 12-APR-1993; 93US-00057952.
PR 04-FEB-1994; 94US-00191708.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
DR WPI; 2000-586558/55.
DR N-PSDB; AAA50664.
XX
PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
PT SIV.
XX
PS Example 1; Col 59-70; 39pp; English.
XX
CC The present sequence is that of fusion protein CD4Bgammal comprising the
CC extracellular portion of CD4, which binds to HIV gp120, linked at its C-
CC terminus to the human IgG1 heavy chain. To obtain the fusion protein, DNA
CC encoding CD4 was linked to IgG1 DNA at the BamI site downstream of the
CC hinge region (see AAA50664). Fusion protein CD4Bgammal and a nucleic acid
CC encoding it are claimed. Also claimed are a vector comprising the nucleic
CC acid, and a method of producing the fusion protein in secreted form using
CC a transformed host cell. The fusion protein may further comprise a
CC therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
CC can be administered to an animal (including humans) for treatment of HIV
CC or SIV infection, and can also be used in assays for HIV or SIV, imaging
CC and tissue stains. IgG1 fusion proteins such as CD4Bgammal provide both
CC complement-mediated and cell-mediated immunity
XX

```

```

SQ Sequence 616 AA;
Query Match 90.0%; Score 3071; DB 3; Length 616;
Best Local Similarity 96.0%; Pred. No. 1e-157;
Matches 599; Conservative 0; Mismatches 15; Indels 10; Gaps 2;

QY 1 MNRGVPFRHLVLVLTALPAATQGNKYVLGKGGTVEITCTASQKSIQFMKNSNQIK 60
DB 1 MNRGVPFRHLVLVLTALPAATQGNKYVLGKGGTVEITCTASQKSIQFMKNSNQIK 60
QY 61 ILGNQSPFLTKGPSKLNBRADSRSLMDQGNFLLIKNIKIEDSPYICEVEDQKEEVL 120
DB 61 ILGNQSPFLTKGPSKLNBRADSRSLMDQGNFLLIKNIKIEDSPYICEVEDQKEEVL 120
QY 121 LVFGILTANSDTHLQGGSLTTLTSPGSSPSVQCRSPKCKNIQCCCKTIVSQLELDSDSC 180
DB 121 LVFGILTANSDTHLQGGSLTTLTSPGSSPSVQCRSPKCKNIQCCCKTIVSQLELDSDSC 180
QY 181 TWTCIVLQKQKVEFIDIVLAFQKASSIVYKKEGQVEFAPPLAFVTEKLTGSGELMW 240
DB 181 TWTCIVLQKQKVEFIDIVLAFQKASSIVYKKEGQVEFAPPLAFVTEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSXKVTODPKLQMGKULPLHLTLPQALPOYAGSGLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSXKVTODPKLQMGKULPLHLTLPQALPOYAGSGLTLA 300
QY 301 LEAKTGKLEHVEVNLVYMRATQLOKRLTCGVMGPTSPKMLSLKLENKEKAVSKREKPVNV 360
DB 301 LEAKTGKLEHVEVNLVYMRATQLOKRLTCGVMGPTSPKMLSLKLENKEKAVSKREKPVNV 360
QY 361 LNPEAGMOCCLSDSGQVLLSESNIKVLTWSTPVPDPAPEPSCDKTHTCPELLGSPSYF 420
DB 361 LNPEAGMOCCLSDSGQVLLSESNIKVLTWSTPVPDPAPEPSCDKTHTCPELLGSPSYF 420
QY 421 LFPPKPKDTLMTSRPEVTCVVVDVSHEDPEVKFMVYVDGVENNAKTPREBOINSTYR 480
DB 421 LFPPKPKDTLMTSRPEVTCVVVDVSHEDPEVKFMVYVDGVENNAKTPREBOINSTYR 480
QY 481 VSVLTIVLHODWLNCKEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 540
DB 471 VSVLTIVLHODWLNCKEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKN 530
QY 541 QVSLTCLVKGFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDSKRMQGN 600
DB 531 QVSLTCLVKGFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDSKRMQGN 590
QY 601 VFSCSVMEHALHNHYTKSLSLSP 624
DB 591 VFSCSVMEHALHNHYTKSLSLSP 614

RESULT 13
AAR26531
ID AAR26531 standard; protein; 534 AA.
XX
AC AAR26531;
XX
DT 25-MAR-2003 (revised)
DT 28-JAN-1993 (first entry)
XX
DE Sequence of CD4-IgG1 chimeric heavy chain heterotrimer.
XX
KM CD4-gamma 1 chimeric heavy chain homodimer; expression vector; HIV;
KM therapy; diagnostic agent; inhibition.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 205..302
FT Region /label=CH1
FT Region 303..317
FT Region /label=hinge
FT Region 318..427

```

```

FT      Region      /label= CH2
FT      428..534
FT      /label= CH3
XX      WO9213559-A1.
XX      20-AUG-1992.
XX      10-FEB-1992; 92WO-US001152.
XX      08-FEB-1991; 91US-00654205.
XX      (PROG-) PROGENICS PHARM INC.
XX      Beaudry GA, Maddon PJ;
XX      WPI, 1992-299758/36.
XX      N-PSDB; AAQ27831.
XX      CD4-gamma 1 chimeric heavy chain homo-dimer and its expression vector -
XX      for preventing and treating HIV infection useful as a diagnostic agent.
XX      Example; Fig 4; 88pp; English.
XX      The human CD4 cDNA is excised from the plasmid pSP6T4 and cloned into
XX      M13mp18. In order to excise a fragment containing the CH1 exon of the
XX      human gamma 1 heavy chain gene, the plasmid pBR gamma 1 is digested with
XX      SacII, and the SacII sites are then made flush using T4 DNA polymerase.
XX      The fragment containing the CH1 exon is then purified and ligated to the
XX      M13mp18 (CD4) vector. Oligonucleotide-mediated site-directed mutagenesis
XX      is then performed to juxtapose the CD4 and CH1 sequences in frame. The
XX      CD4-CH1 chimeric gene is then linearized and ligated to the pCt1-Pc1 DNA
XX      fragment of the plasmid pBR gamma 1 containing the hinge, CH2, and CH3
XX      exons of the human gamma 1 chain gene designated CD4-1g1HC-PRC0V (ATCC
XX      75192). (Updated on 25-MAR-2003 to correct PN field.)
XX      Sequence 534 AA;

Query Match      64.8%; Score 2212; DB 2; Length 534;
Best Local Similarity 72.1%; Pred. No. 1.9e-11;
Matches 457; Conservative 19; Mismatches 48; Indels 110; Gaps 10;

QY      1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKGGDTVLTCTASQKSIQFHKMKNQIX 60
DB      1 MNRGVPRHLLVLTQALLPAATQGNKVVLGKGGDTVLTCTASQKSIQFHKMKNQIX 60
QY      61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEVQL 120
DB      61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKNKIEDSDTYICEVEDQKEVQL 120
QY      121 LVFGLTANSDPHLLQGGSLTTLTBSPPSSPSVQCRSRGKNIQGGKTLSTVSOLELDNSG 180
DB      121 LVFGLTANSDPHLLQGGSLTTLTBSPPSSPSVQCRSRGKNIQGGKTLSTVSOLELDNSG 180
QY      181 TWCTVLQNOQKVEKIDIVLAFOKASSIVYKKEGEQVEFSPLAFTVEKLTSGELMW 240
DB      181 TWCTVLQNOQKVEKIDIVLAFOKASSIVYKKEGEQVEFSPLAFTVEKLTSGELMW 240
QY      241 QAEBASSSKSWITPDILKNKEVSVKRVTDPRKQMGKLLPLHLTLPOALPOYAG---SGNL 297
DB      241 QAEBASSSKSWITPDILKNKEVSVKRVTDPRKQMGKLLPLHLTLPOALPOYAG---SGNL 297
QY      216 ----APSSKS-----TSGTAAAGCIVKDYFPEPVTVSMNSGALTSQVH 255
DB      216 ----APSSKS-----TSGTAAAGCIVKDYFPEPVTVSMNSGALTSQVH 255
QY      298 TLALBAKTKGLHGEVNLVVMRATQI-QKNLTCEWGPSPKLMSTLKENKEAKYSKREK 356
DB      298 TLALBAKTKGLHGEVNLVVMRATQI-QKNLTCEWGPSPKLMSTLKENKEAKYSKREK 356
QY      357 PVMVLPNPAAGMWCCLSDSGVLLSENIKVLPTWSTPVCAPAPKPKSCDKTHTC-----P 411
DB      357 PVMVLPNPAAGMWCCLSDSGVLLSENIKVLPTWSTPVCAPAPKPKSCDKTHTC-----P 411
QY      301 KV-----EPKSDKTKHTHTPPCAP 319
DB      301 KV-----EPKSDKTKHTHTPPCAP 319
QY      412 ELLOGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVEVNAKTKPR 471
DB      412 ELLOGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVEVNAKTKPR 471
QY      320 ELLOGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVEVNAKTKPR 379
DB      320 ELLOGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFMWYVDGVEVNAKTKPR 379

```

```

QY      472 EEOYNSTVRVSVLTVLHODPLNGEKYCKVSNKALPAPIEKTISSKAGOPREPOVYTL 531
DB      472 EEOYNSTVRVSVLTVLHODPLNGEKYCKVSNKALPAPIEKTISSKAGOPREPOVYTL 531
QY      380 EEOYNSTVRVSVLTVLHODPLNGEKYCKVSNKALPAPIEKTISSKAGOPREPOVYTL 439
DB      380 EEOYNSTVRVSVLTVLHODPLNGEKYCKVSNKALPAPIEKTISSKAGOPREPOVYTL 439
QY      532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPTLSDGSPFLYSKLT 591
DB      532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPTLSDGSPFLYSKLT 591
QY      440 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPTLSDGSPFLYSKLT 499
DB      440 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPTLSDGSPFLYSKLT 499
QY      592 DKSRWQGNVPSGVMEHALNHTTQKSLSPG 625
DB      592 DKSRWQGNVPSGVMEHALNHTTQKSLSPG 625
QY      500 DKSRWQGNVPSGVMEHALNHTTQKSLSPG 533
DB      500 DKSRWQGNVPSGVMEHALNHTTQKSLSPG 533

RESULT 14
AAR26530
ID      AAR26530 standard; protein; 435 AA.
XX      AAR26530;
AC      AAR26530;
XX      25-MAR-2003 (revised)
DT      28-JAN-1993 (first entry)
XX      Sequence of one chain of a CD4-gamma 1 chimeric heavy chain homodimer.
DE      CD4-gamma 1 chimeric heavy chain homodimer; expression vector; HIV;
XX      therapy; diagnostic agent; inhibition.
XX      Synthetic.
OS      Key
XX      Location/Qualifiers
XX      Region
XX      /label= CD4
XX      /note= "1.25 = preregion"
XX      /label= hinge
XX      /label= hinge
XX      /label= CH2
XX      /label= CH2
XX      /label= CH3
XX      Region
XX      W09213559-A1.
XX      20-AUG-1992.
XX      10-FEB-1992; 92WO-US001152.
XX      08-FEB-1991; 91US-00654205.
XX      (PROG-) PROGENICS PHARM INC.
XX      Beaudry GA, Maddon PJ;
XX      WPI, 1992-299758/36.
XX      N-PSDB; AAQ27830.
XX      CD4-gamma 1 chimeric heavy chain homo-dimer and its expression vector -
XX      for preventing and treating HIV infection useful as a diagnostic agent.
XX      Example; Fig 3; 88pp; English.
XX      Human CD4 cDNA was excised from pSP6T4 and cloned into M13mp18. The 2 kb
XX      PstI/PstI fragment from pBR lambda 1 contg. the human lambda 1 heavy
XX      chain gene (contg. the hinge, CH2 and CH3 exons) was isolated and cloned
XX      into the BAP-treated M13mp18/CD4 vector. To obtain a CD4-lambda 1
XX      chimeric heavy chain gene, oligonucleotide-mediated site-directed
XX      mutagenesis was performed to juxtapose the CD4 and lambda 1 heavy chain
XX      DNA sequences, ligating the CD4 sequence in frame to the hinge exon. The
XX      DNA was then cloned into pCDNA-1 to produce CD4-1g1-PCDNA1 (ATCC 40951).
XX      (Updated on 25-MAR-2003 to correct PN field.)
XX      Sequence 435 AA;

```

Query Match 62.7%; Score 2140; DB 2; Length 435;  
 Best Local Similarity 68.0%; Pred. No. 1.2e-107;  
 Matches 428; Conservative 0; Mismatches 1; Indels 200; Gaps 2;

QY 2 NRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIKI 61  
 DB 1 NRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIKI 60

QY 62 LGNGGSLTKGSKNDRADSRSLMDQGNPFLIIKNLKIETSDPTVCEVEDQKEVQL 121  
 DB 61 LGNGGSLTKGSKNDRADSRSLMDQGNPFLIIKNLKIETSDPTVCEVEDQKEVQL 120

QY 122 VFGILTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 181  
 DB 121 VFGILTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 180

QY 182 WTCTVLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 241  
 DB 181 WTCTVLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240

QY 242 AERASSKSWITFDLKNKEVSKRVTDPKLQMGKLLPHLTLPOALPOYAGSGLTLAL 301  
 DB 204 ----- 203

QY 302 EAKTGKLGHOEVNLVVMARATQLQKVLTCFVWGPTSPKMLSLKLENKAKVSKREKPVWL 361  
 DB 204 ----- 203

QY 362 NPEAGMWQCLLSDSGOVLLESNIKVLPTWSTPVPCEPAPEPKSCDKHTTC-----PELLGG 416  
 DB 204 -----EPKSCDKHTTCPCPAPELLGG 225

QY 417 PSVFLFPKPKDITLMISTPEVTCVVVDVSHEDPEVKFNNYVGVGVHNAKTKPREQYN 476  
 DB 226 PSVFLFPKPKDITLMISTPEVTCVVVDVSHEDPEVKFNNYVGVGVHNAKTKPREQYN 285

QY 477 STYVAVSLTVLHODMNGKSKCKVSKNKLPAIETKISAKQGPPEVYTLTPSRDE 536  
 DB 286 STYVAVSLTVLHODMNGKSKCKVSKNKLPAIETKISAKQGPPEVYTLTPSRDE 345

QY 537 LTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGFFLYSKLTVDKSRW 596  
 DB 346 LTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGFFLYSKLTVDKSRW 405

QY 597 QGQNVFSCSVMEHLAHNHYTOKSLSLSPG 625  
 DB 406 QGQNVFSCSVMEHLAHNHYTOKSLSLSPG 434

RESULT 15  
 AAR26783 standard; protein; 530 AA.  
 AC AAR26783;  
 DT 24-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 06-FEB-1993 (first entry)  
 DE CD4-IgG2 chimeric heavy chain.  
 KW homodimer; soluble CD4; T cell receptor; CD4 antigen; high recovery;  
 KM chimeric; increased serum half life; HIV infection; AIDS; ss.  
 OS Homo sapiens.  
 OS Chimeric.  
 XX  
 FT Key location/Qualifiers  
 FT Domain 1..205  
 FT Domain /label= CD4 domain  
 FT Domain 206..302  
 FT Domain /label= CH1 domain  
 FT Domain 303..312

FT Domain /label= hinge domain  
 FT 313..423  
 FT /label= CH2 domain  
 FT Domain 424..530  
 FT /label= CH3 domain

PN MO9213947-A1.  
 PD 20-ANG-1992.  
 PF 10-FEB-1992; 92WO-US001143.  
 PR 08-FEB-1991; 91US-00653684.  
 PA (PROG-) PROGENICS PHARM INC.  
 PI Beaudry GA, Madden PJ;  
 DR WPI; 1992-300034/36.  
 DR N-PSDB; AAQ28089.  
 PT CD4-gamma-2 and CD4-IgG2 chimera(s) and expression vectors - for  
 PT treatment, prevention and diagnosis of HIV infection.  
 PS Claim 15; Fig 4; 90pp; English.

XX This sequence represents a CD4-IgG2 chimeric heavy chain heterotrimer  
 CC it was produced by expression of the coding mutagenized cDNA (produced as  
 CC described in AAQ28089) in Dhr-CHO cells. The protein is efficiently  
 CC assembled intracellularly and effectively secreted from mammalian cells  
 CC pref. CHO, COS, or myeloma cells as a heterotrimer, enabling high  
 CC recovery and purification from the medium of cells expressing it. It  
 CC possesses increased serum half-life and has increased avidity for HIV cf.  
 CC heavy chain dimers. It can inhibit HIV infection of CD4+ cells and block  
 CC the spread of HIV infection within a patient. Attachment to a detectable  
 CC marker makes it useful in an assay for HIV or SIV infection, and it can  
 CC also be linked to toxins, eg diphtheria, Pseudomonas exotoxin A (domains  
 CC I or II) or the deglycosylated A-chain of ricin. (Updated on 25-MAR-2003  
 CC to correct PN field.) (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 530 AA;  
 SQ

Query Match 62.0%; Score 2116; DB 2; Length 530;  
 Best Local Similarity 68.9%; Pred. No. 2.8e-106;  
 Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

QY 1 MNRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIK 60  
 DB 1 MNRGVFFRHLILVQLALPPATQGNKVVLGKGGDTVELTCTASQKSIQFHMKNSNQIK 60

QY 61 ILNGGSLTKGSKNDRADSRSLMDQGNPFLIIKNLKIETSDPTVCEVEDQKEVQL 120  
 DB 61 ILNGGSLTKGSKNDRADSRSLMDQGNPFLIIKNLKIETSDPTVCEVEDQKEVQL 120

QY 121 LVEGLTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 180  
 DB 121 LVEGLTANSDTHLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGT 180

QY 181 TWCTVLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240  
 DB 181 TWCTVLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPLAFTVEKLTGSGELMW 240

QY 241 QAEARASSKSWITFDLKNKEVSKRVTDPKLQMGKLLPHLTLPOALPOYAG---SGNL 297  
 DB 217 -----PCSRSTSTALGLVMDYFPEPTVSNVNSALTSGVH 255

QY 298 TLAEAKTGKLGHOEVNLVVMARATOL--QKVLTCFVWGPTSPKMLSLKENKAKVSKREK 356  
 DB 256 TFEAVLQSSGLYSLSSVTVYPPSSNFGTQYTCNV-----DHK 292

QY 357 PVAVLNPEAGMWQCLLSDSGOVLLESNIKVLPTWSTPVP-----PCPAPEPKSCDKHTTC 411  
 DB 293 P-----SNTKVDKTVKREKCVCEPCPCPAP----- 317

QY 412 ELTGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVNAKTKPR 471  
 CC : |||||  
 CC 318 --VAGSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVNAKTKPR 375  
 CC : |||||  
 QY 472 EEOYNSTRVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLR 531  
 CC : |||||  
 DB 376 EEOFNSTRFRVSVLTVVHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLR 435  
 CC : |||||  
 QY 532 PSRDELTKNOVSLTCLVKGYFIPSDIAVEWESNQEPENNYKTTTPVLDSDGSFPLYSLTLY 591  
 CC : |||||  
 DB 436 PSREEMTKNOVSLTCLVKGYFIPSDIAVEWESNQEPENNYKTTTPVLDSDGSFPLYSLTLY 495  
 CC : |||||  
 QY 592 DKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 625  
 DB 496 DKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 529  
 CC : |||||  
 RESULT 16  
 ID AAY85080 standard; protein; 530 AA.  
 XX AAY85080;  
 DT 19-JUN-2000 (first entry)  
 XX  
 DE CD4-IgG2 chimeric heterotetramer heavy chain amino acid sequence.  
 XX  
 KW CD4-IgG2 chimeric heavy chain heterotetramer; immunocombinate; treatment;  
 KM cycloxic radionuclide; cell surface glycoprotein; prevent; infection;  
 KM cellular immune response interaction mediator; HIV interaction; staging;  
 KM prognosis; envelope glycoprotein burden; human.  
 OS Homo sapiens.  
 XX  
 PN US6034223-A.  
 XX  
 PD 07-MAR-2000.  
 XX  
 PF 07-JUN-1995; 95US-00477460.  
 XX  
 PR 07-AUG-1992; 92US-00927931.  
 PR 06-AUG-1993; 93MO-US007422.  
 PR 03-FEB-1995; 95US-00379516.  
 XX  
 PA (PROC-) PROGENICS PHARM INC.  
 XX  
 PI Allaway GP, Madden PJ;  
 DR WPI; 2000-269507/23.  
 DR N-PSDB; AAZ98856.  
 XX  
 PT New immunocombinate, used to treat, prevent or image human immune  
 PT deficiency virus infection, comprises radionuclide attached to  
 PT heterotetramer of CD4-immunoglobulin chimeras.  
 XX  
 PS Disclosure; Fig 4; 58pp; English.  
 XX  
 CC This sequence represents the CD4-IgG2 chimeric heavy chain amino acid  
 CC sequence from the CD4-IgG2 chimeric heterotetramer. The invention relates  
 CC to an immunocombinate comprising a cytotoxic radionuclide and a  
 CC heterotetramer of two heavy chains and two light chains. The cytotoxic  
 CC radionuclide is linked to either the heavy chains or the light chains, or  
 CC to all four chains, directly or through a bifunctional chelator. Both  
 CC heavy chains are chimeric CD4-Ig (immunoglobulin) G2 chains encoded by  
 CC vector CD4-IgG2HC-PRC6V (ATCC 75193) and both light chains are chimeric  
 CC CD4-kappa chains encoded by vector CD4-kLC-PRC6V (ATCC 75194). CD4 is a  
 CC non-polymorphic cell surface glycoprotein that is expressed on the  
 CC surface of helper T lymphocytes, cells of the monocyte/macrophage lineage  
 CC and dendritic cells. CD4 associates with major histocompatibility complex  
 CC (MHC) class II molecules on the surface of antigen presenting cells to  
 CC mediate efficient cellular immune response interactions. In humans CD4 is  
 CC the target of interaction with the human immunodeficiency virus HIV. The

CC immunocombinate is used to kill cells infected with HIV, and for treating  
 CC or preventing infection. It is also used for imaging HIV-infected tissues  
 CC (for staging or prognosis of infection, and for assessing efficacy of  
 CC treatments). The immunocombinate is also used to determine the HIV  
 CC envelope glycoprotein burden, once determined, this information is used  
 CC in the staging and prognosis of HIV infected patients. The  
 CC immunocombinate should be active against all strains of HIV (since the  
 CC CD4-gp120 interaction is essential for infection). The heterotetramers  
 CC are assembled intracellularly and secreted efficiently from mammalian  
 CC cells, allowing high recovery and purification from the culture medium.  
 CC They have longer half-life in serum and greater avidity than heavy chain  
 CC dimers  
 XX  
 SQ Sequence 530 AA;  
 Query Match 62.0%; Score 2116; DB 3; Length 530;  
 Best Local Similarity 68.9%; Pred. No. 2.8e-106;  
 Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;  
 QY 1 MNRGVPFRHLVLQALLPATQGNKVVYLGKKGDTVELTCTASQKKSIOFHKNSNQIK 60  
 DB 1 MNRGVPFRHLVLQALLPATQGNKVVYLGKKGDTVELTCTASQKKSIOFHKNSNQIK 60  
 QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLITKNLKIEDSDPYICEVEDQKEVQL 120  
 DB 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLITKNLKIEDSDPYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLQGSQSLTLTSPSPGSSPVQCSPPKGNIOGKTLTSSOLEIYDQSG 180  
 DB 121 LVFGLTANSDTHLQGSQSLTLTSPSPGSSPVQCSPPKGNIOGKTLTSSOLEIYDQSG 180  
 QY 181 TWTCTVLQNKQKVEFIDIVLAFQKASSTLVYKKEEQVEFSPPLAFYVEKLTGSGELMW 240  
 DB 181 TWTCTVLQNKQKVEFIDIVLAFQKASSTLVYKKEEQVEFSPPLAFYVEKLTGSGELMW 240  
 QY 181 TWTCTVLQNKQKVEFIDIVLAFQKASSTLVYKKEEQVEFSPPLAFYVEKLTGSGELMW 240  
 DB 181 TWTCTVLQNKQKVEFIDIVLAFQKASSTLVYKKEEQVEFSPPLAFYVEKLTGSGELMW 240  
 QY 241 QAERASSKSWITFDLNKKEVSVKRYTODPKLQMGKPLPLHLTLPOALPYAG---SGNL 297  
 DB 241 QAERASSKSWITFDLNKKEVSVKRYTODPKLQMGKPLPLHLTLPOALPYAG---SGNL 297  
 QY 217 -----PCSRSTSESTALGCLVVDYFPEPTVSNAGSLTSGVH 255  
 DB 217 -----PCSRSTSESTALGCLVVDYFPEPTVSNAGSLTSGVH 255  
 QY 298 TLALAKTKGLHOENVLVWRATQL-OKNLTCEVMGPTSKMLSLKENKAVKSKREK 356  
 DB 298 TLALAKTKGLHOENVLVWRATQL-OKNLTCEVMGPTSKMLSLKENKAVKSKREK 356  
 QY 256 TFPAYVLSGSLYSLSVTVTPSSNFGTQTYTCNV-----DHK 292  
 DB 256 TFPAYVLSGSLYSLSVTVTPSSNFGTQTYTCNV-----DHK 292  
 QY 357 PVWVNLNPEAGMOCCLSDSQVLBSNITKYLPIWSTFPV-----PCPAPRKSCKDKHTCP 411  
 DB 357 PVWVNLNPEAGMOCCLSDSQVLBSNITKYLPIWSTFPV-----PCPAPRKSCKDKHTCP 411  
 QY 293 P-----SNTKVDKVEKCCVCEPCPAP----- 317  
 DB 293 P-----SNTKVDKVEKCCVCEPCPAP----- 317  
 QY 412 ELTGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVNAKTKPR 471  
 DB 412 ELTGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVNAKTKPR 471  
 QY 318 --VAGSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVNAKTKPR 375  
 DB 318 --VAGSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVNAKTKPR 375  
 QY 472 EEOYNSTRVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLR 531  
 DB 472 EEOYNSTRVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLR 531  
 QY 376 EEOFNSTRFRVSVLTVVHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLR 435  
 DB 376 EEOFNSTRFRVSVLTVVHODMLNGKEYKCKVSNKALPAPIEKTISKAKGPREPQVYTLR 435  
 QY 532 PSRDELTKNOVSLTCLVKGYFIPSDIAVEWESNQEPENNYKTTTPVLDSDGSFPLYSLTLY 591  
 DB 532 PSRDELTKNOVSLTCLVKGYFIPSDIAVEWESNQEPENNYKTTTPVLDSDGSFPLYSLTLY 591  
 QY 592 DKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 625  
 DB 592 DKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 625  
 QY 496 DKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 529  
 DB 496 DKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 529  
 RESULT 17  
 ID AAB67323 standard; protein; 530 AA.  
 XX AAB67323;  
 AC AAB67323;  
 XX  
 DT 23-APR-2001 (first entry)  
 XX  
 DE CD4-IgG2 chimeric heavy chain protein.





```

QY 1 MNRGVPFRHLLVLTQLALLPAPATGKNVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLLVLTQLALLPAPATGKNVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLIQQGSLTTLTLESPSSPSVQCRSPRGKNTIQGKTLSTVQLELDQSG 180
DB 121 LVFGLTANSPTHLIQQGSLTTLTLESPSSPSVQCRSPRGKNTIQGKTLSTVQLELDQSG 180
QY 181 TMTCTVLQNGKVEFEKIDIVLAFQKASSIYKKEGEQVEFSFPLAFTVEKLTGSGELMW 240
DB 181 TMTCTVLQNGKVEFEKIDIVLAF--AST-----KGPV---FPLA----- 216
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPRKIQNGKLLPLHLTLPQALPOYAG--SGNL 297
DB 217 -----PCSRSTSESTALGCLVXQYFPEPVTVSNMNSGALTSQVH 255
QY 298 TLAEAKTGKLGHOENVLVMAATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 256 TFPVAVLQSSGLYSLSVTVTPSSNFGTQTYTCNV-----DHK 292
QY 357 PWWVLNPEAGMWQCLSDSGVLLSESNIKVLPWSTPV-----PCPAPEPKSCDKHTCP 411
DB 293 P-----SNTKVKTYERKCCVCEPCPAP----- 317
QY 412 ELGQPSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFMVYDVGVEVNAKTKPR 471
DB 318 --VAGPSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVQFMVYDVGVEVNAKTKPR 375
QY 472 EEQYNSTFRVSVLTVLHODWLNGKEYCKVSNKALPAPIEKITISKAKGQPREQVYTLR 531
DB 376 EEQGNSTFRVSVLTVLHODWLNGKEYCKVSNKGLPAPIEKITISKAKGQPREQVYTLR 435
QY 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFPLYSKLTIV 591
DB 436 PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFPLYSKLTIV 495
QY 592 DKSRMQQGNVSCSVMEHALNHYTQKSLSPG 625
DB 496 DKSRMQQGNVSCSVMEHALNHYTQKSLSPG 529

RESULT 19
ABG71123 ID ABG71123 standard; protein; 530 AA.
XX
XX ABG71123;
XX
XX 17-JAN-2003 (first entry)
XX
XX CD4-immunoglobulin G2 (IgG2) chimeric heterotetramer.
XX
XX CD4; immunoglobulin G2; Ig gamma2; human immunodeficiency virus-1; HIV-1;
XX mutant; muclein.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 1..25
XX FT /label= signal_peptide
XX FT 26..530
XX FT /note= "Mature CD4-IgG2 chimeric heterotetramer"
XX
XX US6451313-B1.
XX
XX 17-SEP-2002.
XX
XX 07-JUN-1995; 95US-00484661.

```

```

XX 08-FEB-1991; 91US-00651684.
PR 10-FEB-1992; 92MO-US001143.
PR 08-DEC-1992; 92US-00960440.
XX
XX (PROG-) PROGENICS PHARM INC.
PI Maddon PJ, Beaudry GA;
DR N-PSDB; ABS55721.
DR WPI; 2003-038273/03.
XX
XX Novel CD4-immunoglobulin G2 chimeric heterotetramer neutralizes human
PT immunodeficiency virus-1 with two heavy and light chains encoded by
PT expression vectors designated CD4-IgG2HC-prcCMV and CD4-KLC-prcCMV,
PT respectively.
XX
XX Claim 1; Fig 4A-H; 54p; English.
XX
XX The invention describes a purified CD4-immunoglobulin (Ig)G2 chimeric
CC heterotetramer (I) that neutralizes human immunodeficiency virus-1 (HIV-
CC 1) having two heavy chains encoded by an expression vector designated CD4
CC -IgG2HC-prcCMV, and two light chains encoded by expression vector
CC designated CD4-KLC-prcCMV. (I) and a composition (II) comprising (I) or
CC (I) linked to a toxin, are useful for inhibiting HIV infection of a CD4
CC cell, and preventing a subject being infected with HIV by blocking the
CC spread of HIV infection. This is the amino acid sequence of the CD4-
CC immunoglobulin G2 chimeric heavy chain chimeric heterotetramer useful in
CC inhibiting HIV infection
XX
XX Sequence 530 AA:
SO
Query Match 62.0%; Score 216; DB 6; Length 530;
Best Local Similarity 68.9%; Pred. No. 2.8e-106;
Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;
QY 1 MNRGVPFRHLLVLTQLALLPAPATGKNVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
DB 1 MNRGVPFRHLLVLTQLALLPAPATGKNVVLGKGGDTVELTCTASQKSIQPHMKNNOIK 60
QY 61 ILGNQGSFLLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLLTKGPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLIQQGSLTTLTLESPSSPSVQCRSPRGKNTIQGKTLSTVQLELDQSG 180
DB 121 LVFGLTANSPTHLIQQGSLTTLTLESPSSPSVQCRSPRGKNTIQGKTLSTVQLELDQSG 180
QY 181 TMTCTVLQNGKVEFEKIDIVLAFQKASSIYKKEGEQVEFSFPLAFTVEKLTGSGELMW 240
DB 181 TMTCTVLQNGKVEFEKIDIVLAF--AST-----KGPV---FPLA----- 216
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPRKIQNGKLLPLHLTLPQALPOYAG--SGNL 297
DB 217 -----PCSRSTSESTALGCLVXQYFPEPVTVSNMNSGALTSQVH 255
QY 298 TLAEAKTGKLGHOENVLVMAATQL-QKNLTCEVWGPTSPKMLSLKENKAVSKREK 356
DB 256 TFPVAVLQSSGLYSLSVTVTPSSNFGTQTYTCNV-----DHK 292
QY 357 PWWVLNPEAGMWQCLSDSGVLLSESNIKVLPWSTPV-----PCPAPEPKSCDKHTCP 411
DB 293 P-----SNTKVKTYERKCCVCEPCPAP----- 317
QY 412 ELGQPSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFMVYDVGVEVNAKTKPR 471
DB 318 --VAGPSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVQFMVYDVGVEVNAKTKPR 375
QY 472 EEQYNSTFRVSVLTVLHODWLNGKEYCKVSNKALPAPIEKITISKAKGQPREQVYTLR 531
DB 376 EEQGNSTFRVSVLTVLHODWLNGKEYCKVSNKGLPAPIEKITISKAKGQPREQVYTLR 435
QY 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFPLYSKLTIV 591

```

```
DB 436 PSSEBTKQVSLTCLVKGFPSPDIAVEMESNQEPENNYKTPPMLDSGSPFLYSKLTIV 495
QY 592 DKSRMOQGNVSCSVHMEALHNHYTOKSLSLSPG 625
DB 496 DKSRMOQGNVSCSVHMEALHNHYTOKSLSLSPG 529

RESULT 20
AAE37574
ID AAE37574 standard; protein; 450 AA.
XX
AC AAE37574;
XX
DT 27-AUG-2003 (first entry)
XX
DE Human D1D2-Ig alphacp fusion protein.
XX
KW Human; CD4; cluster of differentiation factor 4; immunoglobulin; Ig;
KW human immunodeficiency virus; gene therapy; vaccine; HIV-1 infection; D1;
KW D2; alpha tailpiece; alphacp; fusion protein.
XX
OS Homo sapiens.
XX
PN MO2003040311-A2.
XX
PD 15-MAY-2003.
XX
PF 24-OCT-2002; 2002MO-US034393.
XX
PR 25-OCT-2001; 2001US-0346231P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Archos J, Cicala C, Fauci AS;
XX
DR WPI; 2003-441545/41.
XX
DR N-PSDB; AAD29113.
XX
PT New CD4 polypeptide ligated at its C-terminus with a portion of an
PT immunoglobulin, useful for preparing a composition for treating or
PT preventing HIV-1 infection.
XX
PS Example 1; Page 47; 100pp; English.
XX
XX
CC The invention relates to a CD4 (cluster of differentiation factor 4)
CC polypeptide ligated at its C-terminus with a portion of an immunoglobulin
CC (Ig) comprising a hinge region and a constant domain of a mammalian Ig
CC heavy chain. The polypeptide comprises a tailpiece from the C-terminus of
CC the heavy chain of an IgA or IgM antibody. Polypeptides of the invention
CC are useful for preparing a composition for treating or preventing human
CC immunodeficiency virus (HIV)-1 infection. The invention is useful in gene
CC therapy and also in the preparation of vaccines. The present sequence is
CC a fusion protein which comprises a human IgA alpha tailpiece (alphacp), a
CC human IgG constant region comprising a hinge, a CH2 and CH3 region and a
CC human CD4 D1D2 domain
XX
SQ Sequence 450 AA;

Query Match 61.7%; Score 2107.5; DB 6; Length 450;
Best Local Similarity 65.7%; Pred. No. 6,9e-106;
Matches 426; Conservative 2; Mismatches 5; Indels 215; Gaps 3;

QY 1 MNRGVPFRHLLLVQLALLPAATQGNKRVLGKKGDVLELTCTASQKKSIQFHWKNSNQIK 60
DB 1 MNRGVPFRHLLLVQLALLPAATQGNKRVLGKKGDVLELTCTASQKKSIQFHWKNSNQIK 60
QY 61 ILNQGSLFKGSKNDRAISRSLMDQGNFPLITNKLIEPDSPTICVEDEQKEVOL 120
DB 61 ILNQGSLFKGSKNDRAISRSLMDQGNFPLITNKLIEPDSPTICVEDEQKEVOL 120
QY 121 LVFGLTANSPTHLLOQGSFLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSPTHLLOQGSFLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
```

```
DB 121 LVFGLTANSPTHLLOQGSFLTLESPPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNKQKVEERKIDIVLAFQKASSIVYKKEGEQVEFPPLAFYBKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEERKIDIVLAFQKASSIVYKKEGEQVEFPPLAFYBKLTGSGELMW 240
QY 241 QABRASSKSWITTFDLKNEVSVKRVTDQPKLQMGKPLPLHLPLQALPQVAGSGLTLA 300
DB 241 QABRASSKSWITTFDLKNEVSVKRVTDQPKLQMGKPLPLHLPLQALPQVAGSGLTLA 300
QY 204 ----- 203
QY 301 LEAKTKLHQEVNLVVMRATOLQKNLTCEWGPSPKMLSLKENKAKVSRKXPVW 360
DB 301 LEAKTKLHQEVNLVVMRATOLQKNLTCEWGPSPKMLSLKENKAKVSRKXPVW 360
QY 204 ----- 203
QY 361 LNPEAGMOCLLSDSGQVILBESNIKVLPTWSTPVCAPAPKSCDKTTC-----PELIG 415
DB 361 LNPEAGMOCLLSDSGQVILBESNIKVLPTWSTPVCAPAPKSCDKTTC-----PELIG 415
QY 416 GPSVFLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNNYVDGVEVNAKTKPREEQY 475
DB 416 GPSVFLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNNYVDGVEVNAKTKPREEQY 475
QY 222 GPSVFLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNNYVDGVEVNAKTKPREEQY 281
DB 222 GPSVFLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNNYVDGVEVNAKTKPREEQY 281
QY 476 NSTYRVSVTLVLDQWLNKGEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRD 535
DB 476 NSTYRVSVTLVLDQWLNKGEYKCKVSNKALPAPIEKTISKAKQPREPOVYTLPPSRD 535
QY 536 ELTKQVSLTCLVKGFPSPDIAVEMESNQEPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595
DB 536 ELTKQVSLTCLVKGFPSPDIAVEMESNQEPENNYKTPPVLDSDGSFFLYSKLTVDKSR 595
QY 342 ELTKQVSLTCLVKGFPSPDIAVEMESNQEPENNYKTPPVLDSDGSFFLYSKLTVDKSR 401
DB 342 ELTKQVSLTCLVKGFPSPDIAVEMESNQEPENNYKTPPVLDSDGSFFLYSKLTVDKSR 401
QY 596 MOQGNVSCSVHMEALHNHYTOKSLSLSPG-----LQLDQETC 632
DB 596 MOQGNVSCSVHMEALHNHYTOKSLSLSPG-----LQLDQETC 632
QY 402 MOQGNVSCSVHMEALHNHYTOKSLSLSGKETHVAVSVMAEVDQTC 449
DB 402 MOQGNVSCSVHMEALHNHYTOKSLSLSGKETHVAVSVMAEVDQTC 449

RESULT 21
AAE37574
ID AAE37574 standard; protein; 530 AA.
XX
AC AAE37574;
XX
XX
XX 25-MAR-2003 (revised)
XX
XX 08-AUG-1994 (first entry)
XX
DE CD4-IgG2 chimeric heavy chain.
XX
KW CD4; gamma; heavy chain; chimeric; chimaeric; immunocjugate; HIV;
KW human immunodeficiency virus; radionuclide; toxin; therapy; treatment;
KW imaging; detection; targeting; immunoglobulin; IgG.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH FT 1..204
FH FT /label= CD4 Region.
FT FT 205..302
FT FT /label= CH1 Region.
FT FT 303..314
FT FT /label= Hinge Region.
FT FT 315..423
FT FT /label= CH2 Region.
FT FT 424..530
FT FT /label= CH3 Region.
XX
XX MO9403191-A1.
XX
XX 17-FEB-1994.
XX
XX 06-AUG-1993; 93MO-US007422.
XX
XX 07-AUG-1992; 92US-00927931.
XX
XX (PROG-) PROGENICS PHARM INC.
```

```

XX Allaway GP, Maddon PJ;
PI WPI: 1994-065392/08.
DR N-PSDB; MA055751.
XX Non-peptidyl toxin or radionuclide and CD4-gamma 2 or CD4-IgG2
PT immuno:conjugates - used to kill HIV-infected cells and to image and
PT stage HIV infection.
XX
PS Disclosure; Fig 4; 142pp; English.
XX
CC A tetramer comprising 2 IgG2 heavy chains or two CD4-IgG2 chimeric heavy
CC chains and two kappa light chains or CD4-kappa light chains (AAR4680)
CC linked to a non-peptidyl toxin or a gamma radiation-emitting radionuclide
CC of low to moderate cytotoxicity. The resulting immunocjugate comprising
CC the toxin can be used to kill HIV infected cells and to treat HIV
CC infected subjects to reduce the population of HIV infected cells. It can
CC also be used to reduce the likelihood of infection. The immunoconjugate
CC comprising the radionuclide can be used to image HIV infected tissue, to
CC calculate the stage of HIV infection or the efficacy of an anti-HIV
CC treatment using the imaging technique and for determining the prognosis
CC of an HIV infected subject. (updated on 25-Mar-2003 to correct PN field.)
XX
SQ Sequence 530 AA;

Query Match 61.7%; Score 2106; DB 2; Length 530;
Best Local Similarity 68.5%; Pred. No. 9.8e-106;
Matches 434; Conservative 28; Mismatches 58; Indels 114; Gaps 10;

1 MNRGVPFRHLLVQLALPPATOGKRVVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
1 MNRGVPFRHLLVQLALPPATOGKRVVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
1 MNRGVPFRHLLVQLALPPATOGKRVVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
121 LVFGLTNSDPHLLQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 180
121 LVFGLTNSDPHLLQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 180
121 LVFGLTNSDPHLLQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 180
181 TWFTVLONOKKVEFKIDIVLAFQKASSIYKKEGQVSRPFLAFTVEKLTGSGELMW 240
181 TWFTVLONOKKVEFKIDIVLAFQKASSIYKKEGQVSRPFLAFTVEKLTGSGELMW 240
181 TWFTVLONOKKVEFKIDIVLAFQKASSIYKKEGQVSRPFLAFTVEKLTGSGELMW 240
181 TWFTVLONOKKVEFKIDIVLAFQKASSIYKKEGQVSRPFLAFTVEKLTGSGELMW 240
241 QAERASSKSWITPDLKNKEVSVKRVYTOPKLCQMKKLPHTLTPQALPOYAG--SGNL 297
241 QAERASSKSWITPDLKNKEVSVKRVYTOPKLCQMKKLPHTLTPQALPOYAG--SGNL 297
241 QAERASSKSWITPDLKNKEVSVKRVYTOPKLCQMKKLPHTLTPQALPOYAG--SGNL 297
217 -----PCSRSTSESTPALGCLVKQYFPEPVVSNMNSGALTSQVH 255
217 -----PCSRSTSESTPALGCLVKQYFPEPVVSNMNSGALTSQVH 255
217 -----PCSRSTSESTPALGCLVKQYFPEPVVSNMNSGALTSQVH 255
298 TLAEATKGLHGVNLVYMRATOL-QKULTCEYWGFTSPKMLSLKENKAVSKREK 356
298 TLAEATKGLHGVNLVYMRATOL-QKULTCEYWGFTSPKMLSLKENKAVSKREK 356
298 TLAEATKGLHGVNLVYMRATOL-QKULTCEYWGFTSPKMLSLKENKAVSKREK 356
256 TFPVILDSGSLYSLSVTVTPSSNFGTQTYCNV-----DHK 292
256 TFPVILDSGSLYSLSVTVTPSSNFGTQTYCNV-----DHK 292
256 TFPVILDSGSLYSLSVTVTPSSNFGTQTYCNV-----DHK 292
357 PWWVLNPDAGMQLDSQGVLESNIKVLPTWSTPV-----PCPAPPEPKSCDKHTCP 411
357 PWWVLNPDAGMQLDSQGVLESNIKVLPTWSTPV-----PCPAPPEPKSCDKHTCP 411
357 PWWVLNPDAGMQLDSQGVLESNIKVLPTWSTPV-----PCPAPPEPKSCDKHTCP 411
293 P-----SNTKVDKTYERKCCVCEPCPAP----- 317
293 P-----SNTKVDKTYERKCCVCEPCPAP----- 317
293 P-----SNTKVDKTYERKCCVCEPCPAP----- 317
412 ELLEGPSVFLFPKPKDTLMISRTPEYTCVVVDVSHEDPEVKFMYDVGVEVNAKTKPR 471
412 ELLEGPSVFLFPKPKDTLMISRTPEYTCVVVDVSHEDPEVKFMYDVGVEVNAKTKPR 471
412 ELLEGPSVFLFPKPKDTLMISRTPEYTCVVVDVSHEDPEVKFMYDVGVEVNAKTKPR 471
318 --VAGPSPVFLFPKPKDTLMISRTPEYTCVVVDVSHEDPEVKFMYDVGVEVNAKTKPR 375
318 --VAGPSPVFLFPKPKDTLMISRTPEYTCVVVDVSHEDPEVKFMYDVGVEVNAKTKPR 375
318 --VAGPSPVFLFPKPKDTLMISRTPEYTCVVVDVSHEDPEVKFMYDVGVEVNAKTKPR 375
472 EBOYNSTYRVVSVLTITVHQMVLNGKEYKCKVSKNKAFLPAPLEKTSKAKGQPREPQVYTLR 531
472 EBOYNSTYRVVSVLTITVHQMVLNGKEYKCKVSKNKAFLPAPLEKTSKAKGQPREPQVYTLR 531
472 EBOYNSTYRVVSVLTITVHQMVLNGKEYKCKVSKNKAFLPAPLEKTSKAKGQPREPQVYTLR 531
376 EBOYNSTYRVVSVLTITVHQMVLNGKEYKCKVSKNKAFLPAPLEKTSKAKGQPREPQVYTLR 435
376 EBOYNSTYRVVSVLTITVHQMVLNGKEYKCKVSKNKAFLPAPLEKTSKAKGQPREPQVYTLR 435
376 EBOYNSTYRVVSVLTITVHQMVLNGKEYKCKVSKNKAFLPAPLEKTSKAKGQPREPQVYTLR 435
532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTIV 591
532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTIV 591
532 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTIV 591
436 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTIV 495
436 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTIV 495
436 PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFPLYSKLTIV 495
592 DKSRMGOQNVFSCSVMEALHNHYTOKSLSPG 625
592 DKSRMGOQNVFSCSVMEALHNHYTOKSLSPG 625
592 DKSRMGOQNVFSCSVMEALHNHYTOKSLSPG 625
496 DKSRMGOQNVFSCSVMEALHNHYTOKSLSPG 529
496 DKSRMGOQNVFSCSVMEALHNHYTOKSLSPG 529
496 DKSRMGOQNVFSCSVMEALHNHYTOKSLSPG 529

```

```

RESULT 22
AAE37576
ID AAE37576 standard; protein; 449 AA.
XX
AC AAE37576;
XX
DT 27-AUG-2003 (first entry)
XX
DE Human FDI2-Ig alpahcp fusion protein variant.
XX
KW Human; CD4; cluster of differentiation factor 4; immunoglobulin; Ig;
KW human immunodeficiency virus; gene therapy; vaccine; HIV-1 infection; D1;
KW D2; alpha tailpiece; alpahcp; fusion protein; muten; variant; mutant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH MISC-difference 218 /note= "Wild type Glu substituted with Pro"
FT MISC-difference 219 /note= "Wild type Leu substituted with Val"
FT MISC-difference 221 /note= "Wild type Gly substituted with Ala"
FT MISC-difference 221 /note= "Wild type Gly substituted with Ala"
XX
PN W02003040311-A2.
XX
PD 15-MAY-2003.
XX
PE 24-OCT-2002; 2002WO-US034393.
XX
PR 25-OCT-2001; 2001US-0346231P.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Athos J, Cicala C, Fauci AS;
XX
DR WPI: 2003-441545/41.
DR N-PSDB; ACC82877.
XX
PT New CD4 polypeptide ligated at its C-terminus with a portion of an
PT immunoglobulin useful for preparing a composition for treating or
PT preventing HIV-1 infection.
XX
PS Example 11; Page 67; 100pp; English.
XX
CC The invention relates to a CD4 (cluster of differentiation factor 4)
CC polypeptide ligated at its C-terminus with a portion of an immunoglobulin
CC (Ig) comprising a hinge region and a constant domain of a mammalian Ig
CC heavy chain. The polypeptide comprises a tailpiece from the C-terminus of
CC the heavy chain of an IgA or IgM antibody. Polypeptides of the invention
CC are useful for preparing a composition for treating or preventing human
CC immunodeficiency virus (HIV)-1 infection. The invention is useful in gene
CC therapy and also in the preparation of vaccines. The present sequence is
CC a fusion protein variant (G218P/L219V/220delA/G221A) which comprises a
CC human IgA alpha tailpiece (alpahcp), a human IgG2 constant region
CC comprising a hinge, a CH2 and CH3 region and a human CD4 DID2 domain.
CC This variant protein is also referred to as mutant F
XX
SQ Sequence 449 AA.

Query Match 61.2%; Score 2088; DB 6; Length 449;
Best Local Similarity 65.2%; Pred. No. 7.7e-105;
Matches 422; Conservative 3; Mismatches 8; Indels 214; Gaps 3;

1 MNRGVPFRHLLVQLALPPATOGKRVVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
1 MNRGVPFRHLLVQLALPPATOGKRVVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
1 MNRGVPFRHLLVQLALPPATOGKRVVLGKGGDTVELTCTASOKKSIOFHMNSNOIK 60
61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
61 ILGNQGSFLTKGSPKLNDRADRSRLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120

```

```

QY 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIQDSG 180
DB 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIQDSG 180
QY 181 TWTCYVLONOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
DB 181 TWTCYVLONOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLOMGKKPLHLTLPLPALPOYAGSGNLTLLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLOMGKKPLHLTLPLPALPOYAGSGNLTLLA 300
QY 204 ----- 203
DB 204 ----- 203
QY 301 LEAKTGKLEHDEVNLYVMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWV 360
DB 301 LEAKTGKLEHDEVNLYVMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWV 360
QY 204 ----- 203
DB 204 ----- 203
QY 361 LNPEAGMOCCLSDSGQVLLLESNIKVLPTWSTVPCBPAPPEKSCDKTHTCPPEL---LGG 416
DB 204 -----SADKTHTCPCBPAPPVAG 221
QY 417 PSVFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTKPRREQYN 476
DB 222 PSVFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTKPRREQYN 281
QY 477 STYRVVSVLTVLHODMLNGKEYCKVSNKALPAPIEKITSKAGQPREPQVYTLPPSRDE 536
DB 282 STYRVVSVLTVLHODMLNGKEYCKVSNKALPAPIEKITSKAGQPREPQVYTLPPSRDE 341
QY 537 LTGNQVSLTCLVGFYPSDIAVEMESNGQPENNYKTPPLVLDSDGSFFLYSKLTVVKSXW 596
DB 342 LTGNQVSLTCLVGFYPSDIAVEMESNGQPENNYKTPPLVLDSDGSFFLYSKLTVVKSXW 401
QY 597 QGQNVFSCSYMHEALHNHYTQKSLSPG-----LQDETC 632
DB 402 QGQNVFSCSYMHEALHNHYTQKSLSPGKPTGVNVSVMVAEVDGTC 448

```

## RESULT 23

AAB19510

ID AAB19510 standard; protein; 481 AA.

XX AAB19510;

XX AC AAB19510;

XX DT 09-JAN-2001 (first entry)

XX DE CD4-IGM fusion protein CH4Pmu.

XX KM CD4; IGM; human; CD4Pmu; fusion protein; immunoglobulin; HIV; SIV; gp120;

XX OS therapy; diagnosis.

XX OS Homo sapiens.

XX XX Key Location/Qualifiers

XX FT Protein 1..395

XX FT Protein /note="CD4 extracellular region"

XX FT Protein 400..481

XX FT Protein /note="IGM heavy chain partial sequence"

XX US6117656-A.

XX PD 12-SEP-2000.

XX PF 07-JUN-1995; 95US-00479353.

XX PR 22-JAN-1988; 88US-00147351.

XX PR 23-JAN-1989; 89US-00299596.

XX PR 09-JUN-1992; 92US-00896781.

XX PR 12-APR-1993; 93US-00057952.

XX PR 04-FEB-1994; 94US-00191708.

XX XX (GEHO ) GEN HOSPITAL CORP.

PI Seed B;

XX WPI; 2000-586558/55.

XX DR N-PSDB; AAA50662.

XX CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or

XX SIV.

XX Example 1; Col 49-60; 39pp; English.

The present sequence is that of fusion protein CD4Pmu comprising the extracellular portion of CD4, which binds to HIV gp120, linked at its C-terminus to the human IGM heavy chain. To obtain the fusion protein, DNA encoding CD4 was linked to IGM DNA at the Pst site upstream of the CH2 region (see AAA50663). Fusion protein CD4Pmu and a nucleic acid encoding it are claimed. Also claimed are a vector comprising the nucleic acid, a method of producing the fusion protein in secreted form using a transformed host cell. The fusion protein may further comprise a therapeutic agent, radiolabel or NMR imaging agent. The fusion protein can be administered to an animal (including humans) for treatment of HIV or SIV infection, and can also be used in assays for HIV or SIV, imaging and tissue stains. IGM fusion proteins such as CD4Pmu provide complement-mediated immunity

SQ Sequence 481 AA;

Query Match 60.7%; Score 2073; DB 3; Length 481;

Best Local Similarity 84.0%; Pred. No. 5,4e-104; Matches 416; Conservative 19; Mismatches 36; Indels 24; Gaps 6;

QY 1 NMRGVFRLHLVLOALLPATQGNKVLGKGGPTVELTCTASQKSIQFMKNSNQIK 60

DB 1 NMRGVFRLHLVLOALLPATQGNKVLGKGGPTVELTCTASQKSIQFMKNSNQIK 60

QY 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120

DB 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEYOL 120

QY 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIQDSG 180

DB 121 LVFGLTANSDTHLLOGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKTLISVQLLEIQDSG 180

QY 181 TWTCYVLONOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240

DB 181 TWTCYVLONOKKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFVTEKLTGSGELMW 240

QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLOMGKKPLHLTLPLPALPOYAGSGNLTLLA 300

DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLOMGKKPLHLTLPLPALPOYAGSGNLTLLA 300

QY 301 LEAKTGKLEHDEVNLYVMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWV 360

DB 301 LEAKTGKLEHDEVNLYVMRATQLOKNLTCEVWGPTSPKMLSLKLENKEAKVSKREKPVWV 360

QY 361 LNPEAGMOCCLSDSGQVLLLESNIKVLPTWSTVPCBPAPPEKSCDKTHTCPPELGG--GPS 418

DB 361 LNPEAGMOCCLSDSGQVLLLESNIKVLPTWSTVPCBPAPPEKSCDKTHTCPPELGG--GPS 418

QY 419 VFLPPPKKDTLMISRPEVTCVVVDVSHEDPEVKFMYVDGVEVHNAKTKPR---E 472

DB 408 VGVFVP-PRDGFPGNPKSKLTCQATGFSR--QIQVSWLRGKQVGSVTTDQVQAEAK 464

QY 473 EQNSTYRVVSVLTIV 487

DB 465 ESGPTTYKVTSTLTI 479

QY 473 EQNSTYRVVSVLTIV 487

DB 465 ESGPTTYKVTSTLTI 479

XX RESULT 24

XX ID AAY59171 standard; protein; 481 AA.

XX AC AAY59171;

XX XX

```

DT 14-MAR-2000 (first entry)
XX
XX CD4-Ig fusion protein CD4Pmu.
DE
XX HIV, extracellular; CD4, gp120; immunoglobulin; Ig; fusion protein;
KM secreted protein; SIV infection; medicament.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX CA1340741-C.
XX
XX 14-SEP-1999.
PD
XX 20-JAN-1989; 89CA-00588749.
PF
XX 20-JAN-1989; 89CA-00588749.
PR
XX 20-JAN-1989; 89CA-00588749.
XX
XX (GEO ) GEN HOSPITAL CORP.
PA
XX
XX Seed B;
PI
XX WPI; 2000-063015/06.
DR N-PSDB; AAZ48204.
DR
XX
XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in
PT the treatment of HIV or simian immunodeficiency virus infections.
XX
XX Example 1; Page 54-60; 89pp; English.
PS
XX The invention provides a fusion gene encoding a fusion protein that
CC comprises an extracellular CD4 DNA sequence or its fragment which binds
CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA
CC sequence of an Ig heavy or light chain, where the DNA sequence encoding
CC the variable region has been replaced with the DNA sequence which encodes
CC extracellular CD4 or its gp120 binding fragment. The fusion protein is
CC capable of being secreted. The fusion proteins are useful for treating
CC HIV or SIV infections in animals, preferably humans. They are also useful
CC for producing medicaments which can be used for treating HIV or SIV
CC infections in humans. The present sequence represents the fusion protein
CC CD4Pmu where the CD4 is linked to human IgG1 at the Fc site upstream of
CC the CH2 region
XX
SQ Sequence 481 AA;
Query Match 60.7%; Score 2073; DB 3; Length 481;
Best Local Similarity 84.0%; Pred. No. 5.4e-104;
Matches 416; Conservative 19; Mismatches 36; Indels 24; Gaps 6;
QY 1 MNRGVPRHLLVQLALPAATQGNKVYGGKGDVVELTCTASOKKSIOFHKNSQIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVYGGKGDVVELTCTASOKKSIOFHKNSQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPPLIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPPLIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDFHLLQGSFLTLTSSPPGSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
DB 121 LVFGLTANSDFHLLQGSFLTLTSSPPGSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
QY 121 LVFGLTANSDFHLLQGSFLTLTSSPPGSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
DB 121 LVFGLTANSDFHLLQGSFLTLTSSPPGSSSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
QY 181 TMTCTVONOKKVEFKIDIVLAFOKASIVYKKEGVEVSFPLAFVEKLTGSGELMW 240
DB 181 TMTCTVONOKKVEFKIDIVLAFOKASIVYKKEGVEVSFPLAFVEKLTGSGELMW 240
QY 241 QAEBASSSKSWITFDLKNKEVSVRKVTQDPFLQMGKLLPLHLTPQALPQYAGSGNTLIA 300
DB 241 QAEBASSSKSWITFDLKNKEVSVRKVTQDPFLQMGKLLPLHLTPQALPQYAGSGNTLIA 300
QY 301 LEATGKLGHOVNLVVRATOLQKQLTCEVWGFTSPKLMSLKLENKEARYSKREKPVWV 360
DB 301 LEATGKLGHOVNLVVRATOLQKQLTCEVWGFTSPKLMSLKLENKEARYSKREKPVWV 360
QY 301 LEATGKLGHOVNLVVRATOLQKQLTCEVWGFTSPKLMSLKLENKEARYSKREKPVWV 360
DB 301 LEATGKLGHOVNLVVRATOLQKQLTCEVWGFTSPKLMSLKLENKEARYSKREKPVWV 360

```

```

QY 361 LNPEAGMOCILSDSGVLLSESNIKVLPWTSTVPCEPAPBPKSCDKTHTCEBELG--GPS 418
DB 361 LNPEAGMOCILSDSGVLLSESNIKVLPWTSTVP-----HADPEVIAELPRK 407
QY 419 VFLFPPKPKDTLMIS-RTPEVTCVVDVSHEDPEVKENMYDVGVENAKTKPR----E 472
DB 408 VSVFVP-PRDGFENPRKSKLICQATGFSR--QIVSWLREGKQVSGVTTDQVQAEAK 464
QY 473 EQNSTRVYVSVLTV 487
DB 465 ESGPTTKVSTLTIT 479

RESULT 25
AA51081
ID AA51081 standard; protein; 481 AA.
XX
XX AA51081;
AC
XX 23-MAR-2000 (first entry)
DT
XX
XX Human fusion protein CD4Pmu.
DE
XX Fusion protein; human; CD4; IgM; immunoglobulin; gp120;
KM anti-human immunodeficiency virus; CD4Pmu.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX US6004781-A.
PN
XX
XX 21-DEC-1999.
PD
XX
XX 04-FEB-1994; 94US-00191708.
PF
XX
XX 22-JAN-1988; 88US-00147351.
PR 23-JAN-1989; 89US-00299596.
PR 09-JUN-1992; 92US-00896781.
PR 12-APR-1993; 93US-00057952.
XX
XX (GEO ) GEN HOSPITAL CORP.
PA
XX
XX Seed B;
PI
XX WPI; 2000-085792/07.
DR N-PSDB; AAZ44064.
DR
XX
XX Fusion protein useful for the treatment of human immunodeficiency virus.
PT
XX
XX Example 1; Col 49-58; 39pp; English.
PS
XX This invention describes a novel nucleic acid (I) encoding a fusion
CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC chain (III). The products of the invention have anti-human
CC immunodeficiency virus (HIV) activity and are capable of binding to
CC gp120. The fusion protein is useful for treating human immunodeficiency
CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence
CC represents the fusion protein CD4Pmu which is constructed from CD4 linked
CC to human IgM upstream of the CH2 region
XX
SQ Sequence 481 AA;
Query Match 60.5%; Score 2067; DB 3; Length 481;
Best Local Similarity 83.8%; Pred. No. 1.1e-103;
Matches 415; Conservative 18; Mismatches 38; Indels 24; Gaps 6;
QY 1 MNRGVPRHLLVQLALPAATQGNKVYGGKGDVVELTCTASOKKSIOFHKNSQIK 60
DB 1 MNRGVPRHLLVQLALPAATQGNKVYGGKGDVVELTCTASOKKSIOFHKNSQIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPPLIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADRSRLMDQGNPPLIKNLKIEDSDTYICEVEDQKEEVOL 120

```

```

Db      61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIIKNLIKIEDSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDFTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDQSG 180
Db      121 LVFGLTANSDFTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDQSG 180
Qy      181 TWTCTVLOQOKVEFKIDIVLAFOKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
Db      181 TWTCTVLOQOKVEFKIDIVLAFOKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
Qy      241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKPLHLTLTPQALPOVAGSGNLTIA 300
Db      241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKPLHLTLTPQALPOVAGSGNLTIA 300
Qy      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEKAVSKREKPVWV 360
Db      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEKAVSKREKPVWV 360
Qy      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPEPKSCDHTCPELLG--GPS 418
Db      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPEPKSCDHTCPELLG--GPS 418
Qy      419 VFLEPPKPKDTLM-ISRTEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPR-----E 472
Db      408 VSVFVP-PRDGFPGCRKSKLICQATGFSR--QIQVSWLRBGKQVGSVTTDQVQAEAK 464
Qy      473 EGYNSTYRVVSULTV 487
Db      465 ESGPTTYKVTSTLTI 479

```

## RESULT 26

AAP93011 standard; protein; 481 AA.

```

ID      AAP93011
AC      AAP93011;
XX
DT      25-MAR-2003 (revised)
DT      03-AUG-1992 (first entry)
XX
DE      Genetic construct which encodes CD4 linked to human Igm at the Pat site
DE      upstream of the CH2 region (fusion protein CD4Pmu).
XX
KM      Fusion protein; immunoglobulin-like molecule; HIV; SIV; therapy;
KM      diagnosis; CD4; gp120; binding fragment; glycoprotein; variable region.
XX
OS      Homo sapiens.
XX
PN      EP325262-A.
XX
PD      26-JUL-1989.
XX
PF      20-JAN-1989; 89EP-00100913.
XX
PR      22-JAN-1988; 88US-00147351.
XX
PA      (GENO ) GEN HOSPITAL CORP.
XX
PI      Seed B;
XX
PI      WPI; 1989-214472/30.
XX
DR      N-PSDB; AAN90359.
XX
PT      Immunoglobulin-CD4 fusion proteins - used for treating HIV or SIV
XX      infections or detecting HIV or SIV in sample.
XX
PS      Example; Table 4, Page 41-47; 68pp; English.
XX
CC      The fusion protein genes of the invention pref. comprises CDNA sequences
CC      which encode CD4 or a fragment which binds gp120 ligated to an expression
CC      plasmid which encodes an antibody in which the variable region of the
CC      gene has been deleted (see WO87-02671). The CD4 portion of the fusion
CC      protein may comprise the complete CD4 sequence, the 370 AA extracellular

```

CC region and the membrane spanning domain, or the extracellular region. The  
 CC Ig heavy chain is pref. from Igm, IgG1 or IgG3. The following are  
 CC specifically claimed: fusion proteins CD4H1ambdA1, CD4Hmu, CD4Pmu,  
 CC CD4IambdA1, and CD4Hmu (No. 67608), PCDAPIambdA (No. 67609) and  
 CC PCDAPIambdA1 (No. 67610). (updated on 25-MAR-2003 to correct PA field.)  
 XX

## SQ Sequence 481 AA;

Query Match 60.5%; Score 2065; DB 1; Length 481;  
 Best Local Similarity 83.8%; Pred. No. 1,4e-103;  
 Matches 415; Conservative 19; Mismatches 37; Indels 24; Gaps 6;

```

Qy      1 NMRGVFPHLLVLQALLPATQGNKVLGKGGDTVELTCTASQKSIQFMKNSNQIK 60
Db      1 NMRGVFPHLLVLQALLPATQGNKVLGKGGDTVELTCTASQKSIQFMKNSNQIK 60
Qy      61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIIKNLIKIEDSDTYICEVEDQKEEVOL 120
Db      61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNPFLIIKNLIKIEDSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDFTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDQSG 180
Db      121 LVFGLTANSDFTHLLQGQSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVSGLEIDQSG 180
Qy      181 TWTCTVLOQOKVEFKIDIVLAFOKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
Db      181 TWTCTVLOQOKVEFKIDIVLAFOKASSIVYKKEGEQVEFSFPLAFVTEKLTGSGELMW 240
Qy      241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKPLHLTLTPQALPOVAGSGNLTIA 300
Db      241 QAERASSSKSWITFDLKNKEVSVKRVYTOPDKLQMGKPLHLTLTPQALPOVAGSGNLTIA 300
Qy      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEKAVSKREKPVWV 360
Db      301 LEAKTGKLGHOEVNLYVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEKAVSKREKPVWV 360
Qy      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPEPKSCDHTCPELLG--GPS 418
Db      361 LNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCEPAPEPKSCDHTCPELLG--GPS 418
Qy      419 VFLEPPKPKDTLMIS-RPEEVTCCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPR-----E 472
Db      408 VSVFVP-PRDGFPGCRKSKLICQATGFSR--QIQVSWLRBGKQVGSVTTDQVQAEAK 464
Qy      473 EGYNSTYRVVSULTV 487
Db      465 ESGPTTYKVTSTLTI 479

```

## RESULT 27

AAR26782 standard; protein; 432 AA.

```

ID      AAR26782
AC      AAR26782;
XX
DT      24-OCT-2003 (revised)
DT      25-MAR-2003 (revised)
DT      06-FEB-1993 (first entry)
XX
DE      CD4-gamma2 chimeric heavy chain homodimer.
XX
KM      homodimer; soluble CD4; T cell receptor; CD4 antigen; high recovery;
KM      chimeric; increased serum half life; HIV infection; AIDS; ss.
XX
OS      Homo sapiens.
XX
OS      Chimeric.
XX
FH      Key
FH      Domain
FT      Domain
FT      Domain
FT      Domain
FT      Domain

```

FT XX /label= CH3  
 XX XX MO9213947-A1.  
 XX XX 20-AUG-1992.  
 XX XX 10-FEB-1992; 92WO-US001143.  
 XX XX 08-FEB-1991; 91US-00653684.  
 XX XX (PROG-) PROGENICS PHARM INC.  
 XX XX Beaudry GA, Maddon PJ;  
 XX XX WPI: 1992-300034/36.  
 DR N-PSDB; AAQ28086.  
 XX XX  
 FT CD4-gamma-2 and CD4-IgG2 chimera(s) and expression vectors - for  
 PT treatment, prevention and diagnosis of HIV infection.  
 PS  
 XX Claim 2; Fig 3; 90pp; English.  
 CC This sequence represents a CD4-gamma2 chimeric heavy chain homodimer. It  
 CC was produced by expression of the coding mutagenised cDNA (produced as  
 CC described in AAQ28086) in Dhfr-CHO cells. The protein is efficiently  
 CC assembled intracellularly and effectively secreted from mammalian cells  
 CC pref. CHO, COS, or myeloma cells as a homodimer, enabling high recovery  
 CC and purification from the medium of cells expressing it. It possesses  
 CC increased serum half-life and has increased avidity for HIV cf. heavy  
 CC chain dimers. It can inhibit HIV infection of CD4+ cells and block the  
 CC spread of HIV infection within a patient. Attachment to a detectable  
 CC marker makes it useful in an assay for HIV or SIV infection, and it can  
 CC also be linked to toxins, eg Diphtheria, Pseudomonas exotoxin A (domains  
 CC I or II) or the deglycosylated A-chain of ricin. (Updated on 25-MAR-2003  
 CC to correct PN field.) (Updated on 24-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 432 AA;  
 Query Match 60.3%; Score 2059; DB 2; Length 432;  
 Best Local Similarity 65.6%; Pred. No. 2.7e-103;  
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;  
 QY 1 MNRGVPRHLLVLLVQLALLPATQGNKVLGKGGDTVELTCTASQKSIQPFHMKNSQIK 60  
 DB 1 MNRGVPRHLLVLLVQLALLPATQGNKVLGKGGDTVELTCTASQKSIQPFHMKNSQIK 60  
 QY 61 ILGNQGSFLLTKGPKSLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEEYQL 120  
 DB 61 ILGNQGSFLLTKGPKSLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVEDQKEEYQL 120  
 QY 121 LVFGLTANSDTHLLQGGSLTTLTSSPPSSPSVQCRSPRGKNIQGGKTLSTVQLDQSG 180  
 DB 121 LVFGLTANSDTHLLQGGSLTTLTSSPPSSPSVQCRSPRGKNIQGGKTLSTVQLDQSG 180  
 QY 181 TWCTCTVLOKQKVEFKIDIVLAFOKASIVYKKEGEQVERSPLAFTVEKLTSSGELMW 240  
 DB 181 TWCTCTVLOKQKVEFKIDIVLAFOKASIVYKKEGEQVERSPLAFTVEKLTSSGELMW 240  
 QY 241 QAERASSKSWITFDLKNKEVSVKRVTDPLQMGKLLPLHLTLPALPOYAGSGNLTLA 300  
 DB 241 QAERASSKSWITFDLKNKEVSVKRVTDPLQMGKLLPLHLTLPALPOYAGSGNLTLA 300  
 QY 301 LEAKTGLHGEVNLVWNRATOLQKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVW 360  
 DB 301 LEAKTGLHGEVNLVWNRATOLQKNTCEVWGPTSPKMLSLKLENKAKVSKREKPVW 360  
 QY 361 LNPAGMWQCLLSDSGVLLSNIKVLPWTGSTVPVCPAPBEKSCDKHTHCEPLLGPSVF 420  
 DB 361 LNPAGMWQCLLSDSGVLLSNIKVLPWTGSTVPVCPAPBEKSCDKHTHCEPLLGPSVF 420  
 QY 421 LFPKPKDTLMIKSTPEVTCVVDVSHEDPEVKENWYVDGVEVNAKTKPREEOYNSTYR 480  
 DB 421 LFPKPKDTLMIKSTPEVTCVVDVSHEDPEVKENWYVDGVEVNAKTKPREEOYNSTYR 480  
 QY 480 LFPKPKDTLMIKSTPEVTCVVDVSHEDPEVKENWYVDGVEVNAKTKPREEOYNSTYR 480  
 DB 480 LFPKPKDTLMIKSTPEVTCVVDVSHEDPEVKENWYVDGVEVNAKTKPREEOYNSTYR 480

QY 481 VVSVLTVLHODMNLGKCKSVKSLPAPIEKTISKAKQPREPQYTLPPSRDELTKN 540  
 DB 287 VVSVLTVHODMNLGKCKSVKSLPAPIEKTISKAKQPREPQYTLPPSRDELTKN 540  
 QY 541 QVSLTCLVKGFPYPSDIAVEMESNGQFPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 600  
 DB 347 QVSLTCLVKGFPYPSDIAVEMESNGQFPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 600  
 QY 601 VFSCSVNHEALHNHYTKSISLSPG 625  
 DB 407 VFSCSVNHEALHNHYTKSISLSPG 431  
 RESULT 28  
 AAR46678  
 ID AAR46678 standard; protein; 432 AA.  
 AC AAR46678;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 08-AUG-1994 (first entry)  
 XX  
 DE CD4-gamma 2 chimeric heavy chain.  
 XX  
 KW CD4; gamma; heavy chain; chimeric; chimeric; immunoconjugate; HIV;  
 KW human immunodeficiency virus; radionuclide; toxin; therapy; treatment;  
 KW imaging; detection; targeting.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..204 /label= CD4 Region.  
 FT Region 205..216 /label= CD4 Region.  
 FT Region 217..325 /label= Hinge Region.  
 FT Region 326..432 /label= CH2 Region.  
 FT Region 433..432 /label= CH3 Region.  
 FT  
 XX  
 XX MO9403191-A1.  
 XX  
 XX 17-FEB-1994.  
 PD  
 XX 06-AUG-1993; 93WO-US007422.  
 PE  
 XX 07-AUG-1992; 92US-00927931.  
 PR  
 XX (PROG-) PROGENICS PHARM INC.  
 PA  
 XX Allway GP, Maddon PJ;  
 PI  
 XX WPI: 1994-065392/08.  
 DR N-PSDB; AAQ57750.  
 DR  
 XX Non-peptidyl toxin or radionuclide and CD4-gamma 2 or CD4-IgG2  
 FT immunoconjugates - used to kill HIV-infected cells and to image and  
 PT stage HIV infection.  
 PS  
 XX Disclosure; Fig 3; 142pp; English.  
 XX  
 CC A CD4-gamma 2 chimeric heavy chain homodimer is linked to a non-peptidyl  
 CC toxin or a gamma radiation-emitting radionuclide of low to moderate  
 CC cytotoxicity. The resulting immunoconjugate comprising the toxin can be  
 CC used to kill HIV infected cells and to treat HIV infected subjects to  
 CC reduce the population of HIV infected cells. It can also be used to  
 CC reduce the likelihood of infection. The immunoconjugate comprising the  
 CC radionuclide can be used to image HIV infected tissue, to calculate the  
 CC stage of HIV infection or the efficacy of an anti-HIV treatment using the  
 CC imaging technique and for determining the prognosis of an HIV infected  
 CC subject. (Updated on 25-MAR-2003 to correct PN field.)



SQ Sequence 432 AA;  
 Query Match 60.3%; Score 2059; DB 2; Length 432;  
 Best Local Similarity 65.6%; Pred. No. 2.7e-103;  
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

QY 1 MNRGVFRRLLLVLTQALLPAATQGNKVVYLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
   |||||
DB 1 MNRGVFRRLLLVLTQALLPAATQGNKVVYLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
   |||||
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
   |||||
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
   |||||
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
   |||||
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
QY 181 TWTCVTLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
   |||||
DB 181 TWTCVTLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
QY 181 TWTCVTLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
   |||||
DB 181 TWTCVTLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
QY 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
   |||||
DB 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
QY 211 -----
   |||||
DB 211 -----
   |||||
QY 301 LEAKTGKLGHOEVNLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKEAKVSKREKPVWV 360
   |||||
DB 301 LEAKTGKLGHOEVNLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKEAKVSKREKPVWV 360
QY 211 -----
   |||||
DB 211 -----
   |||||
QY 361 LNPBAGMOCGLSDSGOVLLSESNIKVLPMTSPVPCPAPKPSCKDHTHCPELLGGPSVF 420
   |||||
DB 361 LNPBAGMOCGLSDSGOVLLSESNIKVLPMTSPVPCPAPKPSCKDHTHCPELLGGPSVF 420
QY 211 -----
   |||||
DB 211 -----
   |||||
QY 421 LFPFKRDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVVDGEVHNAAKTRREOYNSTYR 480
   |||||
DB 421 LFPFKRDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVVDGEVHNAAKTRREOYNSTYR 480
QY 227 LFPFKRDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVVDGEVHNAAKTRREOYNSTYR 286
   |||||
DB 227 LFPFKRDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVVDGEVHNAAKTRREOYNSTYR 286
QY 481 VNSVLTVLHODMNLNGEKYKCKVSNKALPAPIETISKAKQPREPOVYTLPPSRDELTKN 540
   |||||
DB 481 VNSVLTVLHODMNLNGEKYKCKVSNKALPAPIETISKAKQPREPOVYTLPPSRDELTKN 540
QY 287 VNSVLTVLHODMNLNGEKYKCKVSNKALPAPIETISKAKQPREPOVYTLPPSRDELTKN 346
   |||||
DB 287 VNSVLTVLHODMNLNGEKYKCKVSNKALPAPIETISKAKQPREPOVYTLPPSRDELTKN 346
QY 541 QVSLTCLVNGFVPSDIAVWESNGQPENNYKTPPVLDSDGSEFLYSLTVDSKRMQOQN 600
   |||||
DB 541 QVSLTCLVNGFVPSDIAVWESNGQPENNYKTPPVLDSDGSEFLYSLTVDSKRMQOQN 600
QY 347 QVSLTCLVNGFVPSDIAVWESNGQPENNYKTPPVLDSDGSEFLYSLTVDSKRMQOQN 406
   |||||
DB 347 QVSLTCLVNGFVPSDIAVWESNGQPENNYKTPPVLDSDGSEFLYSLTVDSKRMQOQN 406
QY 601 VFSCSVWHEALHNHYTQKSLSLSPG 625
   |||||
DB 601 VFSCSVWHEALHNHYTQKSLSLSPG 625
QY 407 VFSCSVWHEALHNHYTQKSLSLSPG 431
   |||||
DB 407 VFSCSVWHEALHNHYTQKSLSLSPG 431
  
```

RESULT 29  
 AAAY85079  
 ID AAAY85079 standard; protein; 432 AA.  
 AC AAAY85079;  
 DT 19-JUN-2000 (first entry)  
 DE Human CD4-gamma 2 chimeric heavy chain homodimer amino acid sequence.  
 KM CD4-gamma 2 chimeric heavy chain homodimer; immunoglobulin; treatment;  
 KM cytotoxic radionuclide; cell surface glycoprotein; prevent; infection;  
 KM cellular immune response interaction mediator; HIV interaction; staging;  
 KM prognosis; envelope glycoprotein burden; human.  
 OS Homo sapiens.  
 PN US6043223-A.  
 PD 07-MAR-2000.  
 XX 07-JUN-1995; 95US-00477460.

XX 07-AUG-1992; 92US-00927931.  
 PR 06-AUG-1993; 93MO-US007422.  
 PR 03-FEB-1995; 95US-00379516.  
 XX (PROG-) PROGENICS PHARM INC.  
 PA Allaway GP, Madden PJ;  
 PI WPI; 2000-269502/23.  
 DR N-PSDB; AA298855.  
 XX New immunoglobulin, used to treat, prevent or image human immune  
 PT deficiency virus infection, comprises radionuclide attached to  
 PS heterodimer of CD4-immunoglobulin chimeras.  
 XX Disclosure; Fig 3; 58p; English.  
 XX This sequence represents the human CD4-gamma 2 chimeric heavy chain  
 CC homodimer amino acid sequence. The invention relates to an  
 CC immunoglobulin comprising a cytotoxic radionuclide and a heterodimer  
 CC of two heavy chains and two light chains. The cytotoxic radionuclide is  
 CC linked to either the heavy chains or the light chains, or to all four  
 CC chains, directly or through a bifunctional chelator. Both heavy chains  
 CC are chimeric CD4-Ig (immunoglobulin) G2 chains encoded by vector CD4-  
 CC IgG2HC-PRC/MV (ATCC 75193) and both light chains are chimeric CD4-kappa  
 CC chains encoded by vector CD4-kLC-PRC/MV (ATCC 75194). CD4 is a non-  
 CC polymorphic cell surface glycoprotein that is expressed on the surface of  
 CC helper T lymphocytes, cells of the monocyte/macrophage lineage and  
 CC dendritic cells. CD4 associates with major histocompatibility complex  
 CC (MHC) class II molecules on the surface of antigen presenting cells to  
 CC mediate efficient cellular immune response interactions. In humans CD4 is  
 CC the target of interaction with the human immunodeficiency virus HIV. The  
 CC immunoglobulin is used to kill cells infected with HIV, and for treating  
 CC or preventing infection. It is also used for imaging HIV-infected tissues  
 CC (for staging or prognosis of infection, and for assessing efficacy of  
 CC treatments). The immunoglobulin is also used to determine the HIV  
 CC envelope glycoprotein burden, once determined, this information is used  
 CC in the staging and prognosis of HIV infected patients. The  
 CC immunoglobulin should be active against all strains of HIV (since the  
 CC CD4-gp120 interaction is essential for infection). The heterodimers  
 CC are assembled intracellularly and secreted efficiently from mammalian  
 CC cells, allowing high recovery and purification from the culture medium.  
 CC They have longer half-life in serum and greater avidity than heavy chain  
 CC dimers

SQ Sequence 432 AA;  
 Query Match 60.3%; Score 2059; DB 3; Length 432;  
 Best Local Similarity 65.6%; Pred. No. 2.7e-103;  
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

```

QY 1 MNRGVFRRLLLVLTQALLPAATQGNKVVYLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
   |||||
DB 1 MNRGVFRRLLLVLTQALLPAATQGNKVVYLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
   |||||
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
   |||||
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
   |||||
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
   |||||
DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQSGKTLVSQLELDSDG 180
QY 181 TWTCVTLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
   |||||
DB 181 TWTCVTLQNKKEFKIDIVLAFOKASSIVYKKEGQVEFSPFLAFVTEKLTGSGELMW 240
QY 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
   |||||
DB 241 QAEKASSKSWITFDLKNKEVSVKRVYODPKLOMGKKLPLHLTLPOALPOYAGSGNLTLLA 300
QY 211 -----
   |||||
DB 211 -----
   |||||
QY 301 LEAKTGKLGHOEVNLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKEAKVSKREKPVWV 360
   |||||
DB 301 LEAKTGKLGHOEVNLVVMRATQLOKNTLCEVWGPTSPKMLSLKENKEAKVSKREKPVWV 360
  
```

```

Db      211 ----- 210
Qy      361 LNPEAGMOCGLSDSGVLLSNIKVLPTWSTPVPCPAPBPKSCDKHTTCPELLGSPVF 420
Db      211 -----PCCPAPP-----VAGPSVF 226
Qy      421 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDDGVEHNNAKTKPREEQYNSTYR 480
Db      227 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDDGVEHNNAKTKPREEQYNSTYR 286
Qy      481 VVSIVLTIVHODMLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPPVYTLLPPSRDELTKN 540
Db      287 VVSIVLTIVHODMLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPPVYTLLPPSRDELTKN 346
Qy      541 QVSLTCLVKGFPYSDIAVENESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRMWOQN 600
Db      347 QVSLTCLVKGFPYSDIAVENESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRMWOQN 406
Qy      601 VFSCSVNHEALHNHYTOKSLSPG 625
Db      407 VFSCSVNHEALHNHYTOKSLSPG 431

RESULT 30
AAB67322
ID AAB67322 standard; protein; 432 AA.
XX
AC AAB67322;
XX
DT 23-APR-2001 (first entry)
XX
DE CD4-gamma2 chimeric heavy chain homodimer protien.
XX
KM Immunconjugate; chelator; chimeric; HIV; human immunodeficiency virus.
XX
OS Homo sapiens.
XX
PN US6177549-B1.
XX
PD 23-JAN-2001.
XX
PF 10-JUN-1999; 99US-00329916.
XX
PR 07-AUG-1992; 92US-00927931.
PR 06-AUG-1993; 93WO-US007422.
PR 03-FEB-1995; 95US-00379516.
PR 07-JUN-1995; 95US-00477460.
XX
PA (PROG-) PROGENICS PHARM INC.
PI Maddon PJ, Allaway GP;
PI WPI; 2001-158582/16.
XX
DR WPI; 2001-158582/16.
XX
PT Immunconjugate for treating human immunodeficiency virus-infected
PT subject, consists of cytotoxic radionuclide linked to heterotetramer
PT comprising two chimeric CD4-IgG2 heavy chains and two chimeric CD4-
PT kappa light chains.
XX
PS Disclosure; Fig 3; 43pp; English.
XX
CC The present invention relates to an immunconjugate, comprising a
CC cytotoxic radionuclide linked, directly or via a bifunctional chelator,
CC to a heterotetramer having two chimeric CD4-IgG2 heavy chains encoded by
CC an expression vector CD4-IgG2HC-PRCMV and two chimeric CD4-kappa light
CC chains encoded by an expression vector CD4-kLC-PRCMV. The invention is
CC useful for killing human immunodeficiency virus (HIV)-infected cells, for
CC the treatment and prevention of infection with HIV
XX
SQ Sequence 432 AA;
XX
Query Match 60.3%; Score 2059; DB 4; Length 432.

```

```

Best Local Similarity 65.6%; Pred. No. 2,7e-103;
Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

Qy      1 MNRGVPFRHLLVLTQLALLPAATGAKVYLGGKDTVELTCTASQKSIQPHWNSQIK 60
Db      1 MNRGVPFRHLLVLTQLALLPAATGAKVYLGGKDTVELTCTASQKSIQPHWNSQIK 60
Qy      61 ILGNQSFLLTGSPKLNDRADSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120
Db      61 ILGNQSFLLTGSPKLNDRADSRSLMDQGNFLLIKNLKTEDSDTYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLOQSLTTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLOQSLTTLTSPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWCTTVLQONQKVEFKIDIVYLAFOKASSLVYKKEGQVEPSFLATVEKLTGSGELMW 240
Db      181 TWCTTVLQONQKVEFKIDIVYLAFOKASSLVYKKEGQVEPSFLATVEKLTGSGELMW 240
Qy      241 QAERASSKSWITFDLKNKEVSARVYQDPKLGKRLPLHLTLPQALPOYAGSGNLTLA 300
Db      211 ----- 210
Qy      301 LEAKTGKLBQBNLVNRRATOLQKILTCFVWGPTSPKILMLSKLENKAKVSKKEPEVMV 360
Db      211 ----- 210
Qy      361 LNPEAGMOCGLSDSGVLLSNIKVLPTWSTPVPCPAPBPKSCDKHTTCPELLGSPVF 420
Db      211 -----PCCPAPP-----VAGPSVF 226
Qy      421 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDDGVEHNNAKTKPREEQYNSTYR 480
Db      227 LFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDDGVEHNNAKTKPREEQYNSTYR 286
Qy      481 VVSIVLTIVHODMLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPPVYTLLPPSRDELTKN 540
Db      287 VVSIVLTIVHODMLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPPVYTLLPPSRDELTKN 346
Qy      541 QVSLTCLVKGFPYSDIAVENESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRMWOQN 600
Db      347 QVSLTCLVKGFPYSDIAVENESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRMWOQN 406
Qy      601 VFSCSVNHEALHNHYTOKSLSPG 625
Db      407 VFSCSVNHEALHNHYTOKSLSPG 431

RESULT 31
AAB80883
ID AAB80883 standard; protein; 432 AA.
XX
AC AAB80883;
XX
DT 29-MAY-2001 (first entry)
XX
DE Human CD4-gamma2 chimeric heavy chain homodimer.
XX
KM Human; Anti-HIV; CD4-IgG2 chimeric heterotetramer;
KM immunoglobulin gamma 2.
XX
OS Homo sapiens.
XX
PN US6187748-B1.
XX
PD 13-FEB-2001.
XX
PF 07-JUN-1995; 95US-00485372.
XX
PR 08-FEB-1991; 91US-00653684.
PR 10-FEB-1992; 92WO-US001143.
PR 08-DEC-1992; 92US-00960440.
XX

```

PA (PROG-) PROGENICS PHARM INC.  
 XX Maddon PJ, Beaudry GA;  
 PI  
 DR WPI; 2001-264981/27.  
 DR N-PSDB; AAF71829.  
 XX  
 PT Inhibiting human immunodeficiency virus (HIV) infection of a CD4+ cell,  
 PT or treating a subject having CD4+ cells infected with HIV involves using  
 PT CD4-IgG2 chimeric heterotetramer to form a complex with the HIV.  
 XX  
 PS Disclosure; Fig 3; 55pp; English.  
 XX  
 CC The present invention relates to a method for inhibiting infection of a  
 CC CD4+ cell by HIV. The method comprises contacting the HIV with a CD4-IgG2  
 CC chimeric heterotetramer to form a complex with the HIV (CD4 = cluster of  
 CC differentiation 4; IgG2 = immunoglobulin gamma 2). CD4 is a cell surface  
 CC glycoprotein that is expressed primarily on the surface of T cells. In  
 CC man, CD4 is the target of interaction with HIV. The heterotetramer has  
 CC two heavy and two light chains which are encoded by expression vectors  
 CC CD4-IgG2HC-PRCMV (V1) and CD4-KLC-PRCMV (V2), respectively. The method  
 CC is used to inhibit infection of a CD4+ cell by a HIV and to prevent CD4+  
 CC cells of a subject from becoming infected with HIV. The method is also  
 CC useful for treating a subject having CD4+ cells infected with HIV. The  
 CC present sequence is human fusion protein: CD4-gamma2 chimeric heavy chain  
 CC homodimer. This sequence was used in the method of the present invention  
 CC  
 XX  
 XX Sequence 432 AA;  
 SQ  
 Query Match 60.3%; Score 2059; DB 4; Length 432;  
 Best Local Similarity 65.6%; Pred. No. 2,7e-103;  
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;  
 QY 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
 DB 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
 QY 61 IIGNQSPLTKGSPKIDNDRSRSLMDGNFPLIKNKIEDSPDYICGVEQKEEYOL 120  
 DB 61 IIGNQSPLTKGSPKIDNDRSRSLMDGNFPLIKNKIEDSPDYICGVEQKEEYOL 120  
 QY 121 LVEGLTANSRDTHLLOGQSLTTLTSPSPSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180  
 DB 121 LVEGLTANSRDTHLLOGQSLTTLTSPSPSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180  
 QY 121 LVEGLTANSRDTHLLOGQSLTTLTSPSPSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180  
 DB 121 LVEGLTANSRDTHLLOGQSLTTLTSPSPSSPVQCSPPRGKNIQGGKTLSSVQLELDQSG 180  
 QY 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240  
 DB 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240  
 QY 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240  
 DB 181 TWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFVEKLTGSGELMW 240  
 QY 241 QAERASSKSWITFDLKNKREVSVKRVQDPKLOMGKPLHLTLPLQALPOYAGSGNLTLLA 300  
 DB 241 QAERASSKSWITFDLKNKREVSVKRVQDPKLOMGKPLHLTLPLQALPOYAGSGNLTLLA 300  
 QY 211 ----- 210  
 DB 211 ----- 210  
 QY 301 LEAKTGKLEHVEVNLVWRATQLOKNLTCVWGPSTPYMLSLKLENKAVSREKRPVWV 360  
 DB 301 LEAKTGKLEHVEVNLVWRATQLOKNLTCVWGPSTPYMLSLKLENKAVSREKRPVWV 360  
 QY 211 ----- 210  
 DB 211 ----- 210  
 QY 361 LNPBAGMOCLLSDSGVLLSNIKVLPTWSTPVPCEPAPFKSCDKTHTCPBELLGSPVF 420  
 DB 361 LNPBAGMOCLLSDSGVLLSNIKVLPTWSTPVPCEPAPFKSCDKTHTCPBELLGSPVF 420  
 QY 211 -----EC-----PPCPAP-----VAGSVF 226  
 DB 211 -----EC-----PPCPAP-----VAGSVF 226  
 QY 421 LFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRREOYSTVR 480  
 DB 421 LFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRREOYSTVR 480  
 QY 227 LFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRREOYSTVR 286  
 DB 227 LFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKPNWVDVEVHNAAKTRREOYSTVR 286  
 QY 481 VWSVLTVLHODMLNGEKYKCKVSNKALPAIEKTIISAKGQPREPQVYTLPPSRDELTKN 540  
 DB 481 VWSVLTVLHODMLNGEKYKCKVSNKALPAIEKTIISAKGQPREPQVYTLPPSRDELTKN 540  
 QY 287 VWSVLTVLHODMLNGEKYKCKVSNKALPAIEKTIISAKGQPREPQVYTLPPSRDELTKN 346  
 DB 287 VWSVLTVLHODMLNGEKYKCKVSNKALPAIEKTIISAKGQPREPQVYTLPPSRDELTKN 346  
 QY 541 QVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPVILSDSGSFFLYSKLTVDXSRMQQGN 600  
 DB 541 QVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPVILSDSGSFFLYSKLTVDXSRMQQGN 600  
 QY 347 QVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPVILSDSGSFFLYSKLTVDXSRMQQGN 406  
 DB 347 QVSLTCLVKGFPYSDIAVWESNGQPENNYKTPPVILSDSGSFFLYSKLTVDXSRMQQGN 406

QY 601 VPSCSYMEALAHNYTOKSLSPG 625  
 DB 407 VPSCSYMEALAHNYTOKSLSPG 431  
 RESULT 32  
 ABG71122  
 ID ABG71122 standard; protein; 432 AA.  
 XX  
 AC ABG71122;  
 XX  
 DT 17-JAN-2003 (first entry)  
 XX  
 DE CD4-gamma2 chimeric heavy chain of the CD4-IgG2 chimeric protein.  
 XX  
 KW CD4; gamma2 heavy chain; human immunodeficiency virus-1; HIV-1; mutant;  
 KW multien.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..25  
 FT Protein /label= Signal\_peptide  
 FT 26..432  
 FT /note= "Mature CD4-gamma2 chimeric heavy chain of the CD4  
 FT -IgG2 chimeric heterotetramer"  
 XX  
 XX US6451313-B1.  
 XX  
 XX 17-SEP-2002.  
 XX  
 PP 07-JUN-1995; 95US-00484681.  
 XX  
 PR 08-FEB-1991; 91US-00653684.  
 PR 10-FEB-1992; 92WO-US001143.  
 PR 08-DEC-1992; 92US-00960440.  
 XX  
 PA (PROG-) PROGENICS PHARM INC.  
 PI Maddon PJ, Beaudry GA;  
 XX  
 DR WPI; 2003-038273/03.  
 DR N-PSDB; ABS55720.  
 XX  
 PT Novel CD4-immunoglobulin G2 chimeric heterotetramer neutralizes human  
 PT immunodeficiency virus-1 with two heavy and light chains encoded by  
 PT expression vectors designated CD4-IgG2HC-PRCMV and CD4-KLC-PRCMV,  
 PT respectively.  
 XX  
 PS Disclosure; Fig 3A-F; 54pp; English.  
 XX  
 CC The invention describes a purified CD4-immunoglobulin (Ig)G2 chimeric  
 CC heterotetramer (II) that neutralises human immunodeficiency virus-1 (HIV-  
 CC 1) having two heavy chains encoded by an expression vector designated CD4  
 CC -IGG2HC-PRCMV, and two light chains encoded by expression vector  
 CC designated CD4-KLC-PRCMV. (II) and a composition (II) comprising (I) or  
 CC (I) linked to a toxin, are useful for inhibiting HIV infection of a CD4  
 CC cell, and preventing a subject being infected with HIV by blocking the  
 CC spread of HIV infection. This is the amino acid sequence of the CD4-  
 CC gamma2 chimeric heavy chain of the CD4-IgG2 chimeric heterotetramer  
 CC useful in inhibiting HIV infection  
 CC  
 XX  
 XX Sequence 432 AA;  
 SQ  
 Query Match 60.3%; Score 2059; DB 6; Length 432;  
 Best Local Similarity 65.6%; Pred. No. 2,7e-103;  
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;  
 QY 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60  
 DB 1 MNRGVPRHLLVQLALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHMKNSNOIK 60

```

OY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
OY 121 LVFGITANSPTHLLOGSLTLTLSPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
DB 121 LVFGITANSPTHLLOGSLTLTLSPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
OY 181 TWTCTVLONOKKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFVETKLTSGGELMW 240
DB 181 TWTCTVLONOKKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFVETKLTSGGELMW 240
OY 241 QAEKASSKSWITFDLNKKEVSVKRVTDQPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAEKASSKSWITFDLNKKEVSVKRVTDQPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
OY 301 LEAKTKLHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
DB 301 LEAKTKLHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
OY 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVPCPAPPEPKSCDKHTTCEPLGGPSVF 420
DB 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVPCPAPPEPKSCDKHTTCEPLGGPSVF 420
OY 421 LFPKPKDTLMISRPEVTCVVVDVSHEDPEVKFNMVYDGVENHAKTKPREEQYNSTYR 480
DB 421 LFPKPKDTLMISRPEVTCVVVDVSHEDPEVKFNMVYDGVENHAKTKPREEQYNSTYR 480
OY 481 VVSUVTLVHOMLNGKEKCKVSNKALPAPIEKTI SKAKGQPREPQVYTLPPSRDELTKN 540
DB 481 VVSUVTLVHOMLNGKEKCKVSNKALPAPIEKTI SKAKGQPREPQVYTLPPSRDELTKN 540
OY 541 QVSLTCLVKGFPSPDIAEWESNGPENNYKTTPEVLDDSGSFYLSKLTVDKSRMOGN 600
DB 541 QVSLTCLVKGFPSPDIAEWESNGPENNYKTTPEVLDDSGSFYLSKLTVDKSRMOGN 600
OY 601 VFSCSVMEALHNHYTKSLSPG 625
DB 601 VFSCSVMEALHNHYTKSLSPG 625
OY 407 VFSCSVMEALHNHYTKSLSPG 431
DB 407 VFSCSVMEALHNHYTKSLSPG 431

RESULT 33
AAVS1080
ID AAVS1080 standard; protein; 436 AA.
XX
AC AAVS1080;
XX
DT 23-MAR-2000 (first entry)
XX
DE Human fusion protein CD4Mg.
XX
KW Fusion protein; human; CD4; IGM; immunoglobulin; gp120;
KW anti-human immunodeficiency virus; CD4Mg.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN US6004781-A.
XX
PD 21-DEC-1999.
XX
PF 04-FEB-1994; 94US-00191708.
XX
PR 22-JAN-1988; 88US-00147351.
PR 23-JUN-1989; 89US-00299596.
PR 09-JUN-1992; 92US-00896781.
PR 12-APR-1993; 93US-00057952.
XX
PA (GENO ) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
DR WPI; 2000-085792/07.

```

```

DR N-PSDB; AAZ44063.
XX
PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX
PS Example 1; Col 41-50; 39pp; English.
XX
CC This invention describes a novel nucleic acid (I) encoding a fusion
CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC chain (III). The products of the invention have anti-human
CC immunodeficiency virus (HIV) activity and are capable of binding to
CC gp120. The fusion protein is useful for treating human immunodeficiency
CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence
CC represents the fusion protein CD4Mg which is constructed from CD4 linked
CC to human IGM upstream of the CHI region
XX
SQ Sequence 436 AA;
XX
Query Match 59.9%; Score 2046.5; DB 3; Length 436;
Best Local Similarity 90.4%; Pred. No. 1.3e-102;
Matches 405; Conservative 4; Mismatches 16; Indels 23; Gaps 3;

OY 1 MNRGVPFRHLVLVQLALPAPATQGNKVLGKGDVLELTCTASQKSIQPHMNSQIK 60
DB 1 MNRGVPFRHLVLVQLALPAPATQGNKVLGKGDVLELTCTASQKSIQPHMNSQIK 60
OY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVOL 120
OY 121 LVFGITANSPTHLLOGSLTLTLSPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
DB 121 LVFGITANSPTHLLOGSLTLTLSPSSPSVQCRSPRGKNIQGGKTLTSSQLELODSG 180
OY 181 TWTCTVLONOKKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFVETKLTSGGELMW 240
DB 181 TWTCTVLONOKKVEFKIDIVLAFQKASSIYKKKEGQVEFSPLAFVETKLTSGGELMW 240
OY 241 QAEKASSKSWITFDLNKKEVSVKRVTDQPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
DB 241 QAEKASSKSWITFDLNKKEVSVKRVTDQPKLQMGKLPPLHLTLPOALPOYAGSGNLTLA 300
OY 301 LEAKTKLHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
DB 301 LEAKTKLHOEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
OY 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVPCPAPPEPKSCDKHTT-----CPE 412
DB 361 LNPEAGMWQCLSDSGQVLLSNIKVLPTWSTPVHADPECPKRPTRKAKLSTPSARTPG 420
OY 413 LLGGPSVFLFPPKPKDTLMISRPEVTC 440
DB 421 SASAPTLF-----PLVSC 433

RESULT 34
AAVS170
ID AAVS170 standard; protein; 474 AA.
XX
AC AAVS170;
XX
DT 14-MAR-2000 (first entry)
XX
DE CD4-Ig fusion protein CD4Mmu.
XX
KW HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;
KW secreted protein; SIV infection; medicament.
XX
OS Synthetic.
XX
OS Homo sapiens.
XX
PN CA1340741-C.
XX

```

PD 14-SEP-1999.  
 XX 20-JAN-1989; 89CA-00588749.  
 PF 20-JAN-1989; 89CA-00588749.  
 PR 20-JAN-1989; 89CA-00588749.  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA Seed B;  
 PI Seed B;  
 XX WPI; 2000-063015/06.  
 DR N-PSDB; AA248203.  
 XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
 PT the treatment of HIV or simian immunodeficiency virus infections.  
 XX Example 1; Page 47-53; 89pp; English.  
 PS  
 XX The invention provides a fusion gene encoding a fusion protein that  
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 CC the variable region has been replaced with the DNA sequence which encodes  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the fusion protein  
 CC CD4mu where the CD4 is linked to human IgG1 at the Met2 site upstream of  
 CC the CH1 region  
 XX  
 SQ Sequence 474 AA;  
 Query Match 59.8%; Score 2041; DB 3; Length 474;  
 Best Local Similarity 91.0%; Pred. No. 2.8e-102;  
 Matches 406; Conservative 5; Mismatches 31; Indels 4; Gaps 2;

QY 1 MNRGVFRRLLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
 DB 1 MNRGVFRRLLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
 QY 61 ILNGQSFLLTKGFSKLNDRADSRSLMDQGNFLLIKNKIEDSDTYICEVEDQKEVOL 120  
 DB 61 ILNGQSFLLTKGFSKLNDRADSRSLMDQGNFLLIKNKIEDSDTYICEVEDQKEVOL 120  
 QY 121 LVFGLTANSDDTHLLQGSQSLTLTLESPGSSPVQCRSPRKNIQGGKTLSSVQLELDQSG 180  
 DB 121 LVFGLTANSDDTHLLQGSQSLTLTLESPGSSPVQCRSPRKNIQGGKTLSSVQLELDQSG 180  
 QY 121 LVFGLTANSDDTHLLQGSQSLTLTLESPGSSPVQCRSPRKNIQGGKTLSSVQLELDQSG 180  
 DB 121 LVFGLTANSDDTHLLQGSQSLTLTLESPGSSPVQCRSPRKNIQGGKTLSSVQLELDQSG 180  
 QY 181 TWTCTVLQGNQKVEFKIDIVVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240  
 DB 181 TWTCTVLQGNQKVEFKIDIVVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240  
 QY 241 QAERASSSKSMITFDLNGKESVYKRVYQDPKLOMGKLPHTLTPALPOLYAGSGNLTLA 300  
 DB 241 QAERASSSKSMITFDLNGKESVYKRVYQDPKLOMGKLPHTLTPALPOLYAGSGNLTLA 300  
 QY 301 LEAKTGLHQBENVLVNRRATQLOKNTLCEVWGPTSPKMLSLKLENKAVSRKRPVWV 360  
 DB 301 LEAKTGLHQBENVLVNRRATQLOKNTLCEVWGPTSPKMLSLKLENKAVSRKRPVWV 360  
 QY 361 LNPBAGWMOCLSSGOVLLESNIKVLPTWSTPVPKAPKSCDKHTTPELLGGSVF 420  
 DB 361 LNPBAGWMOCLSSGOVLLESNIKVLPTWSTPVPKAPKSCDKHTTPELLGGSVF 420  
 QY 421 LFPKPKDMLMISRTPEVTCVVVDVS 446  
 DB 418 TCEPKPTPKAKLS-TESARTPAEDLS 442

RESULT 35  
 AAR27278  
 ID AAR27278 standard; protein: 532 AA.

XX AAR27278;  
 AC 25-MAR-2003 (revised)  
 DT 28-JUL-1995 (first entry)  
 XX CD4:gamma peptide chimeric protein.  
 DE  
 XX Fusion protein; CD4, extracellular domain; zeta, eta; gamma;  
 XX membrane spanning domain; intracellular domain; type I;  
 XX integral membrane homodimer; TCR; T cell antigen receptor;  
 XX extracellular domain; mouse; human; receptor; chimera;  
 XX HPB-ALL tumour cell line; natural killer cell.  
 XX Homo sapiens.  
 OS  
 XX MO9215322-A1.  
 PN  
 XX 17-SEP-1992.  
 PD  
 XX 06-MAR-1992; 92WO-US001785.  
 PF  
 XX 07-MAR-1991; 91US-00665961.  
 PR  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA Seed B, Romeo C, Kolanus W;  
 XX WPI; 1992-331474/40.  
 DR N-PSDB; AAQ28706.  
 XX  
 PT Therapeutic cells expressing chimeric receptors - directing cellular  
 PT response to an infective agent, useful in treating HIV-1, AIDS  
 PT Pneumocystis carinii infections etc.  
 XX  
 XX Example 2; Page 74-76; 114pp; English.  
 PS  
 XX This sequence represents a fusion protein between the CD4 extracellular  
 CC domain and the gamma protein membrane spanning domain and intracellular  
 CC domain. The FC-receptor-associated gamma chain is expressed in cell  
 CC surface complexes with additional polypeptides, some of which mediate  
 CC ligand recognition, and others which have undefined function. Gamma bears  
 CC a homodimeric structure and overall organisation very similar to that of  
 CC zeta (see also AAQ28704), and is a component of both the mast  
 CC cell/basophil high affinity IgE receptor, FC-epsilon-RI, which consists  
 CC of at least three distinct polypeptide chains and one of the low affinity  
 CC receptors for IgG, represented in mice by FC-gamma-RII-alpha. In the  
 CC production of the CD4 receptor chimera, the gamma cDNA was isolated from  
 CC the HPB-ALL tumour cell line and from human natural killer cells. The  
 CC gamma cDNA was joined to the extracellular domain by engineering a BamHI  
 CC site just upstream of the membrane spanning domain, by a BamHI site  
 CC naturally present a few residues upstream of the membrane spanning  
 CC domain. (Updated on 25-MAR-2003 to correct PN field.)  
 CC  
 XX  
 SQ Sequence 532 AA;  
 Query Match 59.7%; Score 2036.5; DB 2; Length 532;  
 Best Local Similarity 98.3%; Pred. No. 5.5e-102;  
 Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVFRRLLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
 DB 1 MNRGVFRRLLLVQLALPAATQGNKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
 QY 61 ILNGQSFLLTKGFSKLNDRADSRSLMDQGNFLLIKNKIEDSDTYICEVEDQKEVOL 120  
 DB 61 ILNGQSFLLTKGFSKLNDRADSRSLMDQGNFLLIKNKIEDSDTYICEVEDQKEVOL 120  
 QY 121 LVFGLTANSDDTHLLQGSQSLTLTLESPGSSPVQCRSPRKNIQGGKTLSSVQLELDQSG 180  
 DB 121 LVFGLTANSDDTHLLQGSQSLTLTLESPGSSPVQCRSPRKNIQGGKTLSSVQLELDQSG 180  
 QY 181 TWTCTVLQGNQKVEFKIDIVVLAFOKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240

```

Db      ||| 181 TWCTVLOKQKVEFKIDIVLAFQKASSIYKKEGQVERSPFLATFVEKLTSSGELMW 240
Qy      ||| 241 QAERASSSKSWITPDLKNKEVSVKRVYTOPDKLQNGKKLPLHLTLPOALPOYAGSGNLTIA 300
Db      ||| 241 QAERASSSKSWITPDLKNKEVSVKRVYTOPDKLQNGKKLPLHLTLPOALPOYAGSGNLTIA 300
Qy      ||| 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
Db      ||| 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
Qy      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVPKPAPEPKSC 404
Db      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVPKPAPEPKSC 401

RESULT 36
AAR78678 ID AAR78678 standard; protein; 532 AA.
XX
AC AAR78678;
XX
DT 16-APR-1996 (first entry)
XX
DE T-cell receptor eta.
XX
KM Chimeric receptor; CD4; T-cell receptor eta; HIV; cytolysis;
KW human immunodeficiency virus; adoptive immunotherapy.
XX
OS Homo sapiens.
XX
PN MO9521528-A1.
XX
PD 17-AUG-1995.
XX
PF 12-JAN-1995; 95WO-US000454.
XX
PR 14-FEB-1994; 94US-00195395.
PR 02-AUG-1994; 94US-00284391.
XX
PA (GEHO) GEN HOSPITAL CORP.
PI Seed B, Banapour B, Romeo C, Kolanus W;
XX
DR WPI: 1995-292893/38.
DR N-PSDB; AA096124.
XX
PT Target cytolysis of HIV-infected cells - by chimeric CD4 receptor-bearing
PT cells.
XX
PS Example 2; Page 78-79; 118pp; English.
XX
CC Fusion proteins comprising the extracellular domain of CD4 fused to T-
CC cell receptor zeta, gamma or eta (AAR78676-78, respectively) were
CC expressed in CV1 using a vaccine virus vector. These CD4:zeta, CD4:gamma
CC and CD4:eta chimeric receptors mediated cytolysis of targets expressing
CC HIV gp120/41
XX
SQ Sequence 532 AA;

Query Match 59.7%; Score 2036.5; DB 2; Length 532;
Best Local Similarity 98.3%; Pred. No. 5.5e-102;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

```

```

Db      ||| 121 LVFGLTANSPTHLLOQGSFLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      ||| 181 TWCTVLOKQKVEFKIDIVLAFQKASSIYKKEGQVERSPFLATFVEKLTSSGELMW 240
Db      ||| 181 TWCTVLOKQKVEFKIDIVLAFQKASSIYKKEGQVERSPFLATFVEKLTSSGELMW 240
Qy      ||| 241 QAERASSSKSWITPDLKNKEVSVKRVYTOPDKLQNGKKLPLHLTLPOALPOYAGSGNLTIA 300
Db      ||| 241 QAERASSSKSWITPDLKNKEVSVKRVYTOPDKLQNGKKLPLHLTLPOALPOYAGSGNLTIA 300
Qy      ||| 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
Db      ||| 301 LEAKTGKLGHOEVLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYWV 360
Qy      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVPKPAPEPKSC 404
Db      ||| 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVPKPAPEPKSC 401

RESULT 37
AAR89458 ID AAR89458 standard; protein; 532 AA.
XX
AC AAR89458;
XX
DT 26-SEP-1996 (first entry)
XX
DE CD4:eta fusion protein.
XX
KM CD7; transmembrane domain; chimeric receptor; CD5; CD34; CH2; CH3; IgG1;
KW human; CD4; HIV; proteinaceous alpha-helix; T cell; B cell; neutrophil;
KW dendritic cell; therapy; mammal; infection.
XX
OS Synthetic.
XX
PN WO9603883-A1.
XX
PD 15-FEB-1996.
XX
PF 26-JUL-1995; 95WO-US009468.
XX
PR 02-AUG-1994; 94US-00284391.
PR 24-FEB-1995; 95US-00394388.
XX
PA (GEHO) GEN HOSPITAL CORP.
PI Seed B, Banapour B, Romeo C, Kolanus W;
XX
DR WPI: 1996-128034/13.
DR N-PSDB; AAT10803.
XX
PT Membrane-bound chimeric receptor comprising extracellular portion
PT including CD4 fragment - cells expressing receptor can be used for
PT treatment of HIV infection.
XX
PS Example 2; Page 80-81; 134pp; English.
XX
CC AAT10801-T10803 represent membrane bound proteinaceous chimeric receptors
CC of the invention. This sequence represents the CD4:eta chimera. The
CC transmembrane region of the chimeric receptor acts to separate the
CC intracellular and extracellular domains of the chimera, and contains a
CC portion of the CD7 (see AAR89440), CD5 or CD34 transmembrane domains.
CC Alternatively, the extracellular portion of the receptor can be separated
CC from the intracellular domain by the hinge, CH2 and CH3 domains of human
CC IgG1 (see AAR89441). The extracellular portion of the chimeric receptor
CC contains a fragment of CD4 (amino acids 1-394 or 1-200 of the CD4
CC sequence, see AAR89450 and AAR89451) which specifically recognises and
CC binds HIV-infected cells, but does not mediate HIV infection. The
CC extracellular domain of the receptor is separated from the cell membrane
CC by 48 or 72 angstroms, or by one or more proteinaceous alpha-helices. The
CC cells expressing the receptor are preferably T cells, B cells,
CC neutrophils, or dendritic cells. The therapeutic cells expressing the

```

CC chimeric receptor are administered to a mammal to treat HIV infection  
XX  
SQ Sequence 532 AA;  
Query Match 59.7%; Score 2036.5; DB 2; Length 532;  
Best Local Similarity 98.3%; Pred. No. 5.5e-102;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;  
QY 1 MNRGVPFRHLILVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVPFRHLILVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIKNLIKEDSDTYICEVEDQKEEYOL 120  
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIKNLIKEDSDTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSRDTLHLQOGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180  
DB 121 LVFGLTANSRDTLHLQOGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180  
QY 181 TWTCTVLQNGKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
DB 181 TWTCTVLQNGKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
QY 241 QAERASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPALPOYAGSGNLTLLA 300  
DB 241 QAERASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPALPOYAGSGNLTLLA 300  
QY 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360  
DB 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360  
QY 361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTPVAH---DPKLC 401  
DB 361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTPVAH---DPKLC 401  
RESULT 38  
AAW83141  
ID AAW83141 standard; protein; 532 AA.  
XX  
AC AAW83141;  
XX  
DT 03-FEB-1999 (first entry)  
XX  
DE Chimeric receptor containing human eta polypeptide.  
XX  
KM Human; zeta; eta; gamma; membrane-bound chimeric receptor; infection;  
KM tumour; cancer cell; autoimmune-generated cell; T cell receptor; CD3;  
KM CD4; B cell receptor; Fc receptor; pathogen; bacterial; fungal;  
KM protozoan; viral.  
XX  
XX Synthetic.  
OS Homo sapiens.  
OS  
PN US5843728-A.  
XX  
PD 01-DEC-1998.  
XX  
PF 05-APR-1995; 95US-00417495.  
XX  
PR 07-MAR-1991; 91US-00665961.  
PR 06-MAR-1992; 92US-00847566.  
PR 28-FEB-1994; 94US-00203866.  
XX  
PA (GEHO ) GEN HOSPITAL CORP.  
XX  
PI Romeo C, Kolanus W, Seed B;  
XX  
DR WPI; 1999-044582/04.  
DR N-PSDB; AAV70157.  
XX  
PT Membrane-bound chimeric receptors - comprising extracellular portion

PT which recognises and binds a target cell and an intracellular portion of  
PT e.g. a T-cell receptor.  
XX  
PS Claim 11; Col 45-48; 57pp; English.  
XX  
CC The present invention describes DNA encoding a membrane-bound chimeric  
CC receptor comprising: (a) an extracellular portion that specifically  
CC recognises and binds a target cell or a target infective agent; and (b)  
CC an intracellular portion of a T-cell receptor CD3, zeta or eta  
CC polypeptide, a B-cell receptor polypeptide or an Fc receptor polypeptide.  
CC The present sequence represents a chimeric receptor containing the human  
CC eta polypeptide. Cells expressing chimeric receptors of the present  
CC invention can be administered to mammals in order to destroy pathogens  
CC (e.g. bacteria, fungi, protozoa or viruses, especially HIV), cancer cells  
CC or autoimmune-generated cells  
XX  
SQ Sequence 532 AA;  
Query Match 59.7%; Score 2036.5; DB 2; Length 532;  
Best Local Similarity 98.3%; Pred. No. 5.5e-102;  
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;  
QY 1 MNRGVPFRHLILVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVPFRHLILVQLALPAATQGNKVLGKKGDVVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIKNLIKEDSDTYICEVEDQKEEYOL 120  
DB 61 ILGNQSSFLTKGPSKLNDRADSRSLMDQGNPPLIKNLIKEDSDTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSRDTLHLQOGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180  
DB 121 LVFGLTANSRDTLHLQOGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQDSG 180  
QY 181 TWTCTVLQNGKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
DB 181 TWTCTVLQNGKVEFKIDIVLAFQKASSIVYKKEGEVFEFPLAFTVEKLTGSGELMW 240  
QY 241 QAERASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPALPOYAGSGNLTLLA 300  
DB 241 QAERASSSKSWITFDLKNKEVSVRVYTOPDKLQMGKKLPLHLTLPALPOYAGSGNLTLLA 300  
QY 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360  
DB 301 LEAKTGKLEHENVLVVWRATOLQKNLTCEVWGPTSPKMLSLKLENKEAKVSRKRPVWY 360  
QY 361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTPVAH---DPKLC 401  
DB 361 LNPBAGMOCCLSDSGQVLLSENIKVLPTWSTPVAH---DPKLC 401  
RESULT 39  
AAR27276  
ID AAR27276 standard; protein; 575 AA.  
XX  
AC AAR27276;  
XX  
DT 25-MAR-2003 (revised)  
DT 28-JUL-1995 (first entry)  
XX  
DE CD4:zeta peptide chimeric protein.  
XX  
KM Fusion protein; CD4; extracellular domain; zeta; eta; gamma;  
KM membrane spanning domain; intracellular domain; type I;  
KM integral membrane homodimer; TCR; T cell antigen receptor;  
KM extracellular domain; mouse; human; receptor; chimera;  
KM HPB-ALL tumour cell line; natural killer cell.  
XX  
XX Homo sapiens.  
OS  
OS  
XX  
XX Key Location/Qualifiers  
FH 1..399  
FT Protein /notes "CD4 extracellular domain"